Energy Economics 151 (2025) 108922

Contents lists available at ScienceDirect

Energy Economics

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

Dealing with endogeneity in stochastic frontier models: A comparative assessment of estimators☆

Zheng Hou [a,b,*], Joaquim J.S. Ramalho [c], Catarina Roseta-Palma [c]

[a] UECE – Research Unit on Complexity and Economics, Lisbon School of Economics and Management, University of Lisbon, Portugal
[b] Cef.up - Center for Economics and Finance, School of Economics and Management, University of Porto, Portugal
[c] Department of Economics and BRU-Business Research Unit, ISCTE - Instituto Universitário de Lisboa, Portugal

ARTICLE INFO

JEL classification:
C36
D24
L23
O25
Q40

Keywords:
Stochastic Frontier Analysis
Technical efficiency
Endogeneity
Instrumental variables
Energy sector

ABSTRACT

Endogeneity poses a major challenge for Stochastic Frontier Analysis, as input choices may be endogenous to unobserved components of the error term, resulting in biased efficiency estimates. This paper compares leading estimators that address this issue, including control-function estimator (Kutlu, 2010), Generalized Method of Moments (GMM) (Tran and Tsionas, 2013) and copula (Tran and Tsionas, 2015) approaches, as well as the instrumental variable based maximum likelihood estimator (Karakaplan and Kutlu, 2017a,b; Karakaplan, 2022). Monte Carlo simulations reveal distinct bias–variance trade-offs: likelihood-based estimators provide more precise efficiency scores, while GMM and copula can be advantageous in specific contexts. An empirical application to the Portuguese thermal power subsector (2006-2021) shows that accounting for endogeneity significantly alters coefficients and efficiency distributions. These results demonstrate that estimator choice affects policy-relevant indicators such as efficiency scores and determinants of cost performance. Despite data limitations, the study underscores the importance of treating endogeneity and provides methodological guidance for applied efficiency analysis.

1. Introduction

Technical efficiency is the predominant focus of Stochastic Frontier Analysis (SFA) applied to firms in the energy sector. The performance of this sector, in turn, underpins overall economic activity. Historically, in the 20th century, many countries structured their energy sectors around vertically integrated monopolies under governmental supervision, since there were high fixed costs of investment and this organization was seen as the way to create a functioning energy system. However, the imperative of enhancing efficiency became evident over time, as the problems inherent to monopoly control surfaced. Consequently, numerous countries embarked on market reforms to stimulate competition, segmenting generation, transmission, and distribution into distinct subsectors with differing technical and economic attributes. While such reforms generally aim to enhance technical efficiency (Barros, 2008; Ma and Zhao, 2015; Lundin, 2020; Bobde and Tanaka, 2020), their effectiveness may be compromised by various challenges stemming from incomplete deregulation or practical obstacles (Sun and Wu, 2020; Lee and Howard, 2021; Mirza et al., 2021). To facilitate market reforms and promote efficiency improvements, regulatory bodies utilize various benchmarking techniques, such as Data Envelopment Analysis (DEA) and SFA.

With the ongoing energy transition, new challenges emerge to energy market regulators, leading to continuing interest in applied research on benchmark methods. Accurate estimations of firms' efficiency levels are important to policy making in order to correctly incentivize inefficient players to catch up with efficient ones. One issue that awaits resolution is how to deal with potential endogeneity in benchmarking methods. This issue has received attention from recent studies (Kumbhakar et al., 2020; Kuosmanen, 2023). Notably, inputs can be endogenous as firms select them according to specific economic goals. Consequently, the inputs within a production function exhibit correlation with the error term, which includes unobserved firm-specific effects, production risk (statistical noise), and technical inefficiency (Lai and Kumbhakar, 2019). Thus, the use of conventional Stochastic Frontier estimators without considering endogeneity is problematic. Early attempts to address the endogeneity issue in Stochastic Frontier models include Guan et al. (2009), Kutlu (2010) and Kim and Kim (2011). They are followed by subsequent studies: Tran and Tsionas (2013, 2015), Griffiths and Hajargasht (2016) and Amsler et al. (2016,

☆ This article is part of a Special issue entitled: 'Benchmarking' published in Energy Economics.

* Corresponding author.
E-mail addresses: zhenghou@uece.iseg.ulisboa.pt (Z. Hou), Joaquim.Jose.Ramalho@iscte-iul.pt (J.J.S. Ramalho), catarina.roseta@iscte-iul.pt (C. Roseta-Palma).

https://doi.org/10.1016/j.eneco.2025.108922
Received 28 June 2024; Received in revised form 16 September 2025; Accepted 16 September 2025
Available online 23 September 2025

2017), among others. Various methods are used to handle endogeneity in Stochastic Frontier estimation, for example, Bayesian approach (Griffiths and Hajargasht, 2016), nonparametric approach (Prokhorov et al., 2021) and Poisson frontier models (Haschka and Herwartz, 2022). Lai and Kumbhakar (2019) and Kumbhakar et al. (2020) show that endogeneity can be eliminated once input misallocation is accounted for, but the theoretical formulation relies on specific assumptions and still does not directly treat endogeneity in the econometrical estimation. Therefore, the endogeneity problem remains unsolved when evaluating technical efficiency in the energy sector, especially when electricity is the object of study.

A maximum likelihood estimator to correct for endogeneity in Stochastic Frontier models is proposed by Karakaplan and Kutlu (2017a) and Karakaplan (2017). The method initially enables the estimation of cross-sectional models; Karakaplan and Kutlu (2017b) and Karakaplan (2022) further develop the panel-data version of the method. For the time being, this is a comprehensive and practical method which has been applied in studies of various economic fields, including energy economics (e.g., Xu et al., 2022). Notwithstanding, it has not yet been applied to evaluate firm efficiency in the energy sector.

In particular, endogeneity needs to be addressed in benchmark assessments of the sector in order to obtain reliable efficiency measures and avoid misleading policy conclusions. To this end, we evaluate alternative estimators that have been proposed to handle endogeneity in stochastic frontier models: control function estimator (Kutlu, 2010); GMM estimator (Tran and Tsionas, 2013); copula-based estimator (Tran and Tsionas, 2015); instrumental variable (IV) based maximum likelihood estimator (Karakaplan and Kutlu, 2017b; Karakaplan, 2022).[1] In doing so, the paper goes beyond a single application and establishes a comparative benchmark for the use of endogeneity-corrected frontier methods in applied research. A further contribution is the explicit implementation of these estimators (except for the IV-based maximum likelihood estimator *xtsfkk*, already available as a Stata package) in R script, provided in the replication package. This not only enhances transparency and reproducibility but also equips researchers with ready-to-use tools that can be adapted to diverse empirical settings, thereby facilitating the broader adoption of econometric approaches to efficiency analysis.

In the simulations, we find that IV-based maximum likelihood estimators, GMM, and copula methods differ in their bias–variance tradeoffs, with the likelihood approach generally providing more precise efficiency scores, while GMM and copula can be advantageous in specific scenarios. In the empirical application to the Portuguese thermal power subsector, these methodological differences translate into noticeable variation in estimated efficiency distributions and policy-relevant indicators. Together, the findings underscore the importance of explicitly accounting for endogeneity when applying frontier models in energy and related sectors.

The rest of this article is organized as follows. In the next section we review related literature. In Section 3 we describe the methodological formulations of representative approaches correcting for endogeneity in SFA that we select for Monte Carlo simulations and empirical applications. In Section 4 we describe the implementation of Monte Carlo simulations of these representative estimators and compare simulation results. Section 5 describes the empirical model and data used for the empirical application, and presents and discusses the results. Concluding remarks are made in Section 6.

[1] In fact, the control function estimator is also an IV-based maximum likelihood approach, which is stated differently here to avoid confusion.

2. Literature review

Various approaches have been proposed for evaluating technical efficiency, particularly in the energy sector. An early example is the *best practice efficiency measure* by Diewart and Nakamura (1999), based on Farrell (1957). Jamasb and Pollitt (2001) provide an overview of prevalent benchmarking methodologies in the electricity industry, including DEA, Corrected Ordinary Least Squares (COLS), and SFA, along with their primary applications up to that point. Recently, there has been a growing preference for DEA and SFA in assessing technical efficiency. Unlike traditional methods, DEA does not impose specific functional forms on input and output, thereby enabling researchers to concentrate on efficiency concerns (Ma and Zhao, 2015). However, it does not allow direct estimation of the relationship between efficiency and explanatory factors; when it is indeed justified to do so, a two-stage approach is often applied (e.g., Bigerna et al., 2019, 2020, 2022), where extra care is necessary to prevent biased estimations. DEA has found diverse applications in energy economics, as evidenced by recent studies. Gultom (2019) analyze efficiency in the U.S. electricity sector; Navarro-Chávez et al. (2020) examine the Mexican electricity sector; Alizadeh et al. (2020) investigate the Iranian electricity sector; Jindal and Nilakantan (2021) investigate Indian coal-fired power plants; Vesterberg et al. (2021) examine the Swedish electricity distribution; Sánchez-Ortiz et al. (2021) study the Spanish electricity sector; and Nakaishi et al. (2021) look into the environmental efficiency of Chinese coal-fired power plants, among others.

The Stochastic Frontier approach presupposes technical inefficiency, the inability to attain the output frontier given the inputs and prevailing technology. The method involves making distributional assumptions regarding the noise and inefficiency components (Kumbhakar and Tsionas, 2008), facilitating the estimation of how independent variables influence the mean and variance of technical inefficiency. While some Stochastic Frontier techniques do not necessitate distributional assumptions for the noise or inefficiency components, as demonstrated by Kumbhakar and Bernstein (2019), alternatives such as nonlinear squares can also obviate the need for such assumptions, as exemplified by Belotti and Ferrara (2021).

Stochastic Frontier models offer adaptability, allowing customization to tackle a diverse array of issues, thereby more effectively catering to specific research inquiries. Thus, SFA has garnered widespread acceptance in the field of energy economics. It is a versatile tool, suitable for evaluating various aspects such as directed technological change (Hou et al., 2020, 2021), profitability and the long-term viability of energy production options (Lee and Howard, 2021). Llorca et al. (2017) analyze efficiency in the Latin-American transport sector using energy demand functions. Kumbhakar et al. (2020) assess the cost efficiency of Norwegian electricity distribution firms, contributing to the growing body of research employing SFA to address efficiency concerns in the electricity sector. Other studies focusing on technical efficiency in Norwegian electricity distribution include Growitsch et al. (2012), Kumbhakar et al. (2015), Kumbhakar and Lien (2017), Orea et al. (2018), Mydland et al. (2018), Musau et al. (2021), among others. Soroush et al. (2021) investigate the impact of institutional quality on cost efficiency in Italian electricity distribution utilities. Vesterberg et al. (2021) study the efficiency of small-scale electricity and distribution grid in Sweden.

When it comes to electricity generation, Lai and Kumbhakar (2018) introduce a homoscedastic four-component stochastic frontier (H4CSF) model, linking technical inefficiency in production to factors such as the age and capacity of coal-fired power plants. Liu et al. (2019) explore whether environmental variations affect the technical efficiency of Chinese grid utilities. Silva et al. (2019) employ a stochastic frontier approach with maximum entropy estimation to analyze European electricity distribution companies. Peñasco et al. (2019) delve into the influence of policy factors on the efficiency of Spanish solar energy plants.

Early attempts to mitigate endogeneity in SFA adopted IV-based control-function strategies. Kutlu's (2010) two-step "BCIV" estimator projects each potentially endogenous input on external instruments and then inserts the first-stage residuals into a Battese–Coelli likelihood (Battese and Coelli, 1992).[2] Karakaplan and Kutlu (2017a,b), extend this idea to the *sfkk/xtsfkk* estimators, which jointly estimate the frontier and reduced-form equations for the endogenous regressors in one maximum-likelihood step. Guan et al. (2009) and Tran and Tsionas (2013) replace the second-step ML with a one-step, just-identified GMM based on score conditions, preserving consistency under weaker distributional assumptions but still relying on valid instruments.

When suitable instruments are unavailable, researchers have turned to the joint-distribution approach. Tran and Tsionas (2015) embed a Gaussian copula in the composed-error likelihood, obtaining a single-step ML estimator that remains consistent without instruments, although it assumes normal rank dependence. Griffiths and Hajargasht (2016) achieve a similar goal with a Chamberlain–Mundlak control-function device that links firm-level input averages to permanent inefficiency; their Bayesian/ML framework can also accommodate a transient inefficiency term. Amsler et al. (2016, 2017) further generalize copula and IV ideas, allowing both inputs and the scaling (environmental) variables that enter the inefficiency term to be endogenous; they offer parallel IV-GMM and simulated-ML estimators. More recently, Haschka and Herwartz (2022) extend the copula logic to a Poisson frontier for count data, again dispensing with instruments.

A second branch of the literature addresses endogeneity by embedding firms' first-order conditions (FOC) directly in the frontier. Building on the H4CSF framework, Lai and Kumbhakar (2018) and Lien et al. (2018) impose the FOC from maximizing return to the outlay, while Lai and Kumbhakar (2019) and Kumbhakar et al. (2020) include the cost-minimizing FOC and decompose the resulting allocative-inefficiency term into persistent and time-varying parts. Because the FOC error term absorbs the correlation between inputs and the composite error, these models control endogeneity "structurally" rather than econometrically. A complementary line relaxes functional-form assumptions: Prokhorov et al. (2021) propose non-parametric and semi-parametric frontiers with endogenous regressors, combining flexible series or kernel approximations with a copula-based dependence structure. To date, these structural-FOC and flexible-form estimators have been applied less frequently than the IV-based *sfkk/xtsfkk* family, largely because they require stronger behavioral assumptions or heavier computation.

IV-based maximum likelihood estimators remain the most frequently applied econometric cure for endogeneity in frontier work. A practical benchmark is the estimator of Karakaplan and Kutlu (2017a), a maximum-likelihood procedure for cross-sectional data that augments the Battese–Coelli frontier with first-stage reduced forms for each endogenous regressor. Karakaplan (2017) released the Stata command *sfkk*, making the method easy to replicate. The same authors extend the idea to panels in Karakaplan and Kutlu (2017b); the corresponding Stata command, *xtsfkk*, is documented in Karakaplan (2022). Subsequent studies have built on this framework: Kutlu et al. (2019) allow the individual inefficiency term to follow a time-varying latent process, while Kutlu (2022) embeds the IV-frontier in a spatial setting with an endogenous weighting matrix. Together these contributions illustrate how the *sfkk/xtsfkk* family has become the work-horse IV platform for applied SFA. The sfkk and xtsfkk estimators have been applied across a wide range of topics, including biomass energy (Xu and H.H., 2018), school expenditures (Karakaplan and Kutlu, 2019), farmland productivity (Deng et al., 2020), market power in iron ore (Germeshausen et al., 2020), innovation in finger millet (Jerop et al., 2020), rice productivity under climate change (Ojo and Baiyegunhi, 2020), corporate social responsibility (CSR) and efficiency (Binh et al., 2022), economic agglomeration and energy efficiency (Xu et al., 2022), rural bank efficiency (Amanda, 2023), bank lending in China (Fungáčová et al., 2023), transport infrastructure and output (Melo-Becerra and Ramírez-Giraldo, 2023), and coal mining efficiency (Yang and Tsou, 2024). To our knowledge, however, no study has yet used the endogeneity estimator to analyze firm efficiency in the energy sector.

Table 1 provides an overview of the typical approaches to tackle endogeneity in SFA, highlighting essential aspects of each one.

Energy market regulators can utilize benchmarking models to estimate the cost efficiency of firms, enabling the establishment of revenue caps and incentivizing improvements in productivity and efficiency without unnecessary micro-management (Kumbhakar et al., 2020). However, neglecting endogeneity when estimating efficiency scores may lead to incorrect information being used for such policy formulation. This study evaluates the performance of alternative estimators designed to handle endogeneity in stochastic frontier models, combining Monte Carlo simulations with an empirical application to the thermal power subsector of Portugal. By jointly considering simulation evidence and a sector-specific application, the analysis highlights both the methodological trade-offs and the practical implications of dealing with endogeneity, thereby motivating the empirical and simulation frameworks developed in the following sections.

3. Overview on representative approaches addressing endogeneity in SFA

In this section, we introduce some representative approaches for addressing endogeneity in SFA that we will use in Monte Carlo simulation and empirical application. These include: (1) the standard maximum likelihood estimator assuming exogenous inputs, which can be referred to as the "naive estimator"; (2) the two-step IV control-function estimator by Kutlu (2010), often referred to as BCIV, which is a basic and accessible solution for panel data; (3) the score-based GMM estimator by Tran and Tsionas (2013), which offers a one-step estimation using external instruments; (4) the copula-based maximum likelihood estimator developed by Tran and Tsionas (2015), which handles endogeneity without relying on external instruments; and (5) the *xtsfkk* estimator proposed by Karakaplan and Kutlu (2017b) and Karakaplan (2022), which provide a maximum likelihood framework to jointly estimate the frontier and reduced-form equations and are available in Stata. These estimators cover different methodological strategies – including control function, GMM, copula approaches and IV-based ML – and thus provide a well-rounded basis for performance comparison.

For each estimator, we introduce the basic model specification and the key steps[3] for estimating the model.

3.1. The naive estimator

The "naive" model in our analysis corresponds to the original stochastic frontier specification of Aigner et al. (1977) and Meeusen and van den Broeck (1977), extended to panel data without introducing time variation in inefficiency or correcting for endogeneity. The production frontier is given by:

$$y_{it} = \mathbf{X}_{it}'\beta + v_{it} - u_{it}, \tag{1}$$

where y_{it} denotes the output of firm i in period t, $\mathbf{X}_{it}$ is a vector of input variables, and β is a vector of technology parameters to be estimated. The composite error term consists of:

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

[2] For control-function endogeneity corrections in time-varying-parameter models outside the SFA framework, see Kim (2006) and Kim and Kim (2011).

[3] To avoid redundancy, we describe simplified formulations for the approaches. The readers may refer to the original articles for complete formulations.

Table 1
Summary of representative empirical approaches that handle endogeneity in SFA.

Study/Estimator	Data required	External instruments	Inefficiency structure assumed	Identification/Estimation technique
Kutlu (2010) – BCIV	Panel	Yes (first-stage OLS)	Time-varying half-normal	Two-step control-function ⟶ ML
Guan et al. (2009)	Panel	Yes (lags as IVs)	Time-varying	Difference-GMM first stage; ML frontier
Karakaplan and Kutlu (2017a) – *sfkk*	Cross-section	Yes	Firm-specific half-normal (scaled by covariates)	Joint ML on frontier + reduced-form equations
Karakaplan (2022) – *xtsfkk*	Panel	Yes	Same as *sfkk*; firm effects handled	Joint ML (frontier + reduced forms)
Tran and Tsionas (2013)	Cross-section (run year-by-year)	Yes	Time-varying score-GMM	One-step, just-identified
Tran and Tsionas (2015)	Cross-section	No	Time-varying	Instrument-free Gaussian-copula ML
Griffiths and Hajargasht (2016)	Panel	No (Chamberlain –Mundlak CRE)	Permanent (optionally plus transient)	CRE control-function, ML/Bayesian
Amsler et al. (2016, 2017)	Cross-section (panel feasible)	Optional (IV-GMM variant)	Model-specific: permanent or time-varying	IV-GMM or simulated ML with copula
Lai and Kumbhakar (2018); Lien et al. (2018)	Panel	No	Persistent + transient via allocative-inefficiency term	Structural FOC of return-to-outlay; unified ML
Lai and Kumbhakar (2019); Kumbhakar et al. (2020)	Panel	No	Persistent + transient (decomposed)	FOC of cost minimization; structural ML
Kutlu et al. (2019)	Panel	Yes (extends *sfkk*/*xtsfkk*)	Time-varying true individual effects	Expanded ML with state-space inefficiency
Prokhorov et al. (2021)	Cross-section	No	Time-varying	Non- & semi-parametric frontier; copula dependence
Haschka and Herwartz (2022)	Cross-section (count data)	No	Time-varying (Poisson frontier)	Instrument-free copula ML for Poisson SFA

where v_{it} captures statistical noise and other random shocks, while $u_{it} \geq 0$ represents firm-specific technical inefficiency, assumed to follow a half-normal distribution N^+ with zero mean and variance σ_u^2. The variance parameters are reparameterized as:

$$\sigma^2 = \sigma_u^2 + \sigma_v^2, \gamma = \frac{\sigma_u^2}{\sigma^2}, \tag{3}$$

where γ measures the proportion of total variance attributable to inefficiency. The model parameters β, σ_u^2 and σ_v^2 are estimated by maximum likelihood, and firm-specific technical efficiency is predicted following (Jondrow et al., 1982):

$$TE_{it} = E[\exp(-u_{it})|\varepsilon_{it}], \varepsilon_{it} = v_{it} - u_{it}. \tag{4}$$

This naive model serves as the baseline for comparison with alternative estimators that address endogeneity.

3.2. The control function estimator

Kutlu (2010) extends the standard stochastic frontier framework of Aigner et al. (1977) and Meeusen and van den Broeck (1977) to allow for endogenous regressors in the production frontier. The baseline production frontier is:

$$y_{it} = \mathbf{X}_{it}'\beta + \varepsilon_{it} - u_{it}, \tag{5}$$

where y_{it} denotes the output of firm i in period t, $\mathbf{X}_{it}'$ is an $m \times 1$ vector of potentially endogenous regressors, β is a vector of technology parameters, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ is statistical noise, and $u_{it} \geq 0$ represents technical inefficiency, assumed to follow a truncated-normal distribution $N^+(\mu, \sigma_u^2)$. Endogeneity arises when $\mathbf{X}_{it}$ is correlated with the noise term ε_{it}. Kutlu (2010) models $\mathbf{X}_{it}$ as

$$\mathbf{X}_{it} = \mathbf{Z}_{it}'\delta + \mathbf{v}_{it}, \tag{6}$$

where $\mathbf{Z}_{it}$ contains exogenous instruments, and $(v_{it}, \varepsilon_{it})$ are jointly normally distributed with correlation vector ρ. By a Cholesky decomposition of the covariance matrix of $(v_{it}, \varepsilon_{it})'$, the production equation can be rewritten as

$$y_{it} = \mathbf{X}_{it}'\beta + \sigma_\varepsilon \rho' \boldsymbol{\Sigma}_v^{-\frac{1}{2}}(\mathbf{X}_{it} - \mathbf{Z}_{it}'\delta) + w_{it} - u_{it}, \tag{7}$$

where w_{it} is independent of v_{it}. In practice, Kutlu (2010) suggests either a joint maximum likelihood estimation, which incorporates both the frontier equation and the reduced form for the endogenous regressors into a single likelihood, or a computationally simpler two-step procedure. In the two-step approach, Eq. (6) – the reduced form for the endogenous regressors – is first estimated by OLS, and the fitted residuals $\hat{v}_{it}$ are then included in the frontier equation as an additional regressor:

$$y_{it} = \mathbf{X}_{it}'\beta + \rho'\mathbf{v}_{it} + w_{it} - u_{it}. \tag{8}$$

This augmented stochastic frontier model is estimated by maximum likelihood under the usual half-normal inefficiency assumption. Firm-specific technical efficiency is predicted as in Jondrow et al. (1982), conditional on both w_{it} and the estimated endogenous component $\hat{v}_{it}$:

$$TE_{it} = E[\exp(-u_{it})|w_{it}, \hat{v}_{it}]. \tag{9}$$

3.3. The GMM estimator

Tran and Tsionas (2013) propose a one-step GMM estimator for stochastic frontier models with endogenous regressors. The starting point is the frontier equation:

$$y_{it} = \mathbf{Z}_{1,it}'\alpha + \mathbf{X}_{it}'\beta + v_{it} - u_{it}, \tag{10}$$

where $\mathbf{X}_{it}$ is a $p \times 1$ vector of potentially endogenous regressors, $\mathbf{Z}_{1,it}'$ is a $q_1 \times 1$ vector of exogenous regressors, $v_{it} \sim N(0, \sigma_v^2)$ is statistical noise, and $u_{it} \sim N^+(0, \sigma_u^2)$ represents non-negative inefficiency. The endogenous regressors are modeled in a reduced form[4]:

$$\mathbf{X}_{it} = \mathbf{Z}_{2,it}'\delta + \varepsilon_{it}, \tag{11}$$

[4] To avoid confusion with the common annotation γ that represents the share of inefficiency variance in the total composed error variance in SFA, here we use δ instead of γ in Tran and Tsionas (2015).

where $\mathbf{Z}_{2,it} = \mathbf{I}_p \otimes \tilde{\mathbf{Z}}_{2,it}$ contains q_2 strictly exogenous variables ($q_2 \geq p$), and (ϵ_{it}, v_{it}) are jointly normally distributed with correlation vector $\boldsymbol{\rho}$. By applying a Cholesky decomposition to the covariance matrix of $(\tilde{\epsilon}_{it}, v_{it})'$, the frontier equation can be re-expressed as:

$$y_{it} = \mathbf{Z}'_{1,it}\boldsymbol{\alpha} + \mathbf{X}'_{it}\boldsymbol{\beta} + \sigma_v \boldsymbol{\rho}' \boldsymbol{\Omega}_\epsilon^{-\frac{1}{2}} (\mathbf{X}_{it} - \mathbf{Z}'_{2,it}\boldsymbol{\delta}) + \omega_{it} - u_{it}, \tag{12}$$

where $\omega_{it} \sim N(0, (1-\boldsymbol{\rho}'\boldsymbol{\rho})\sigma_v^2)$ is independent of $\tilde{\epsilon}_{it}$. The GMM approach uses the first-order conditions from the correct likelihood as moment conditions, combined with the orthogonality conditions from the reduced-form equation. In the exactly identified case, the parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\delta}$,variance parameters$)'$ are estimated by solving:

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{G}_i(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\delta}) = \mathbf{0}. \tag{13}$$

Following the estimation of parameters, firm-specific efficiency is obtained in line with Jondrow et al. (1982). Let $\hat{\epsilon}_{it}$ denote the corrected residual, $\hat{\sigma}_u$ and $\hat{\sigma}_v$ denote the estimated standard deviations of the inefficiency and noise terms, $\hat{\lambda} = \hat{\sigma}_u/\hat{\sigma}_v$, and $\hat{\sigma} = \sqrt{\hat{\sigma}_u^2 + \hat{\sigma}_v^2}$. Then the technical efficiency of firm i at time t is calculated as

$$TE_{it} = \exp[-\hat{\sigma}(\frac{\varphi(\frac{\hat{\epsilon}_{it}\hat{\lambda}}{\hat{\sigma}})}{1-\Phi(\frac{\hat{\epsilon}_{it}\hat{\lambda}}{\hat{\sigma}})} - \frac{\hat{\epsilon}_{it}\hat{\lambda}}{\hat{\sigma}})], \tag{14}$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal probability density (PDF) and cumulative distribution functions (CDF), respectively.

3.4. The copula-based estimator

The copula-based approach models the dependence between the noise term v_{it} and the inefficiency term u_{it} without relying on the joint normality assumption. The stochastic frontier is specified as:

$$y_{it} = \mathbf{X}'_{it}\boldsymbol{\beta} + v_{it} - u_{it}, \tag{15}$$

where $v_{it} \sim N(0, \sigma_v^2)$ and $u_{it} \sim N^+(\mu, \sigma_u^2)$ are allowed to be statistically dependent. The joint distribution $F_{u,v}(u, v)$ is constructed via a copula function $C_\theta(\cdot, \cdot)$:

$$F_{u,v}(u, v) = C_\theta(F_u(u), F_v(v)), \tag{16}$$

where F_u and F_v are the marginal CDFs of u_{it} and v_{it}, and θ is the copula parameter measuring dependence. In our implementation, we employ the Gaussian copula:

$$C_\theta(s, t) = \Phi_\theta(\Phi^{-1}(s), \Phi^{-1}(t)), \tag{17}$$

where Φ^{-1} is the standard normal quantile function, and Φ_θ is the CDF of a bivariate normal distribution with correlation coefficient θ. The log-likelihood function is obtained from the copula density:

$$l(\boldsymbol{\beta}, \sigma_u, \sigma_v, \mu, \theta) = \sum_{i=1}^{N}\sum_{t=1}^{T} \ln C_\theta(F_u(u_{it}), F_v(v_{it})) + \ln f_u(u_{it}) + \ln f_v(v_{it}), \tag{18}$$

where C_θ is the copula density and f_u, f_v are the marginal densities. Parameters are estimated by maximizing this log-likelihood. Firm-specific technical efficiency is computed as

$$TE_{it} = E[\exp(-u_{it})|v_{it}], \tag{19}$$

where the conditional distribution of u_{it} given v_{it} is derived from the joint copula-based specification.

3.5. The xtsfkk estimator

The estimator proposed by Karakaplan and Kutlu (2017b) and Karakaplan (2022), implemented in Stata as *xtsfkk*, generalizes the control-function approach of Kutlu (2010) to a panel-data framework and unlike the previous estimators, allows for endogeneity in both the frontier and inefficiency equations. The model is based on the following specification:

$$\begin{aligned} y_{it} &= \mathbf{X}'_{y,it}\boldsymbol{\beta} + v_{it} - su_{it}, \\ \mathbf{X}_{it} &= \mathbf{Z}_{it}\boldsymbol{\delta} + \epsilon_{it}, \end{aligned} \tag{20}$$

where y_{it} is output, $\mathbf{X}_{y,it}$ includes both exogenous and endogenous variables in the frontier, v_{it} is the two-sided noise term, u_{it} is the one-sided inefficiency term, and $s = 1$ for production frontiers and $s = -1$ for cost frontiers. The reduced form equation links endogenous regressors $\mathbf{X}_{it}$ to instruments $\mathbf{Z}_{it}$. By applying a Cholesky decomposition of the joint distribution of (ϵ_{it}, v_{it}), the frontier can be rewritten with a bias-correction component:

$$y_{it} = \mathbf{X}'_{y,it}\boldsymbol{\beta} + (\mathbf{X}_{it} - \mathbf{Z}_{it}\boldsymbol{\delta})'\eta + e_{it}, \tag{21}$$

where $(\mathbf{X}_{it} - \mathbf{Z}_{it}\boldsymbol{\delta})'\eta$ captures the correlation between endogenous regressors and the two-sided error, and $e_{it} = w_{it} - su_{it}$ with w_{it} independent from regressors. The inefficiency is modeled as

$$u_{it} = h(\mathbf{X}'_{u,it}\boldsymbol{\varphi}_u)u_i^*, u_i^* \sim N^+(\mu, \sigma_u^2). \tag{22}$$

For each panel unit i, the log-likelihood function is decomposed as

$$\ln L_i = \ln L_{i,y|\mathbf{X}} + \ln L_{i,\mathbf{X}}, \tag{23}$$

where $\ln L_{i,y|\mathbf{X}}$ is the conditional density of output given regressors and the correction term, and $\ln L_{i,\mathbf{X}}$ corresponds to the reduced-form equations for the endogenous regressors. Technical efficiency is then predicted as

$$TE_{it} = \exp\{-h_{it}[\mu_{i*} + \sigma_{i*}\frac{\phi(\mu_{i*}/\sigma_{i*})}{\Phi(\mu_{i*}/\sigma_{i*})}]\}, \tag{24}$$

where $\phi(\cdot)$, $\Phi(\cdot)$ are the standard normal PDF and CDF.

Additionally, endogeneity can be assessed by testing the joint significance of the bias-correction term $(\mathbf{X}_{it} - \mathbf{Z}_{it}\boldsymbol{\delta})\eta$, where η loads the first-stage residuals. The test relies on similar ideas with the standard Durbin–Wu–Hausman test for endogeneity. Under exogeneity these residuals are orthogonal to the composed error, hence $\eta = 0$. Therefore it tests

$$H_0 : \eta = 0 \text{ vs. } H_1 : \eta \neq 0.$$

The test statistic is the Wald chi-square

$$W = \hat{\eta}'[Var(\hat{\eta})]^{-1}\hat{\eta} \overset{a}{\sim} \chi^2(q), \tag{25}$$

with q the number of endogenous regressors. Rejection implies endogeneity.

Similarly, for the models introduced in Sections 3.2–3.4, In the R implementation, endogeneity is tested by a Wald test applied to the model-specific correction parameters.

4. Monte Carlo simulations

This section aims to evaluate the performance of the representative estimators mentioned in the previous section through a controlled Monte Carlo simulation. The objective is twofold: first, to illustrate the magnitude of bias and efficiency loss when endogeneity is ignored in Stochastic Frontier estimation; and second, to compare the performance of representative estimators that have been developed to address endogeneity, and therefore, evaluate their strengths and weaknesses in application.

In each of the above mentioned estimators that address endogeneity in SFA, endogeneity is assumed to be present between only the regressor and the noise term except in the *xtsfkk* estimator. Thus, two DGPs (data-generating processes) are adopted to suit the assumptions of different estimators. One DGP assumes correlation between the regressor(s) and the noise term, while another DGP follows Karakaplan (2022) for the *xtsfkk* estimator, which targets endogeneity between the regressor(s), the inefficiency term and the noise term. For each simulation, we compare the mean bias and Root Mean Squared Error (RMSE)[5] between each estimated coefficient and its true value for assessing the accuracy of estimation. The DGPs and implementation of the estimators are described as follows.

4.1. Endogeneity between the regressor and the noise term

4.1.1. Data-generating process

The DGP follows a standard Cobb–Douglas stochastic frontier specification with additive composite error and endogenous regressors. For each simulation replication, we generate a balanced panel with $N = 100$ cross-sectional units observed over $T = 20$ time periods. The outcome variable y_{it} is generated as:

$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + v_{it} - u_{it}, \tag{26}$$

$$x_{2it} = z_{it} + \varepsilon_{it}, \tag{27}$$

$$v_{it} = \rho\varepsilon_{it} + \sqrt{1-\rho^2}\,v_{it}^{*}, \tag{28}$$

where y_{it} denotes the output for unit i at time t; x_{1it} is an exogenous input; x_{2it} is an endogenous input; z_{it} is an instrumental variable used to construct x_{2it}. The composed error consists of a two-sided noise term $v_{it} \sim N(0, \sigma_v^2)$ and a non-negative inefficiency term $u_{it} = u_i$, with $u_i \sim N^{+}(0, \sigma_u^2)$.[6] The endogeneity arises through correlation between v_{it} and ε_{it}, parameterized by $\rho \in [0, 1]$. The idiosyncratic error terms follow ε_{it}, $v_{it}^{*} \sim N(0, 1)$, independently across units and time. Simulations are performed for varying levels of endogeneity $\rho \in \{0, 0.4, 0.8\}$ with 1000 repetitions each. The true parameter values are set to $\beta_1 = \beta_2 = 0.5$, $\sigma_u^2 = \sigma_v^2 = 1$. Therefore, the inefficiency share $\gamma = \frac{\sigma_u^2}{\sigma_u^2+\sigma_v^2} = 0.5$.

4.1.2. Implementation of the estimators

The naive model. The naive Stochastic Frontier model is estimated using maximum likelihood under the standard composed error specification, where the inefficiency term $u_{it} \sim N^{+}(0, \sigma_u^2)$ is assumed to be time-invariant and independent of the regressors. The model takes the form $y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + v_{it} - u_{it}$, where $v_{it} \sim N(0, \sigma_v^2)$ captures statistical noise. This specification ignores potential endogeneity in the regressors and thus serves as a benchmark. The log-likelihood function is constructed following the standard normal-truncated normal formulation, and parameters are estimated via the BFGS algorithm using the maxLik package in R.

The Battese–Coelli estimator. To account for endogeneity, the Battese–Coelli estimator of Kutlu (2010) is implemented by augmenting the frontier equation with residuals obtained from the first-stage regression of the endogenous regressor on its instruments. In the first stage, x_{2it} is regressed on z_{it}, and the residual $r_{it} = x_{2it} - \hat{x}_{2it}$ is extracted. This residual is then included as an additional regressor in the stochastic frontier model: $y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \delta r_{it} + v_{it} - u_{it}$. The model is estimated by maximum likelihood using the standard composed error specification, where $v_{it} \sim N(0, \sigma_v^2)$ and $u_{it} \sim N^{+}(0, \sigma_u^2)$. By incorporating the control function r_{it}, correlation between x_{2it} and the noise term is captured, allowing consistent estimation without requiring external instruments in the second stage.

The GMM estimator. The GMM estimator follows Tran and Tsionas (2013), with necessary adaptations. It is implemented by constructing a system of orthogonality conditions that account for the endogeneity of regressors within the stochastic frontier framework. First, residuals from the reduced-form equation $x_{2it} = z_{it}\theta + \varepsilon_{it}$ are computed, and these residuals are included in the structural equation as a control function: $y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \delta r_{it} + v_{it} - u_{it}$, where $r_{it} = x_{2it} - \hat{x}_{2it}$. The resulting moment conditions are derived from the score functions of the log-likelihood with respect to the parameters of interest, and include both analytical terms and numerical derivatives with respect to the inefficiency and noise variances. A Q-minimization procedure is then applied to the average moment vector using a BFGS optimizer, yielding consistent estimates under the assumption that the instruments are valid and the inefficiency term $u_{it} = u_i \sim N^{+}(0, \sigma_u^2)$ is independent of the regressors. Asymptotic efficiency is achieved through joint estimation of both structural and reduced-form parameters.

The copula-based estimator. The estimation of the copula-based estimator follows Tran and Tsionas (2015), with necessary adaptations. It is implemented by jointly modeling the dependence between the endogenous regressor and the composed error term using a Gaussian copula function. The stochastic frontier equation is specified as $y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + v_{it} - u_{it}$, with $v_{it} \sim N(0, \sigma_v^2)$ and $u_{it} \sim N^{+}(0, \sigma_u^2)$. The composite error $\varepsilon_{it} = y_{it} - (\beta_1 x_{1it} + \beta_2 x_{2it})$ is transformed into a skew-normal distribution, and its probability integral transform is paired with that of x_{2it} using a Gaussian copula. The log-likelihood is constructed from the joint density implied by the copula and the marginal distributions of x_{2it} and ε_{it}. The parameters are first initialized using a global optimization routine (DEoptim), and then refined using maximum likelihood estimation via a BFGS optimizer. This approach allows the correlation between the endogenous regressor and the composed error to be directly estimated without requiring external instruments.

4.1.3. Simulation results and discussion

The above models are simulated with 1000 repetitions, in each of which using the same generated sample across different models. The simulation results are demonstrated in Table 2.

The naive stochastic frontier estimator, which ignores the endogeneity of regressors, performs well only when the correlation between the regressors and the composed error is absent ($\rho = 0$). In this case, the estimates of both slope coefficients and inefficiency parameters are nearly unbiased. However, as endogeneity increases ($\rho = 0.4$ and 0.8), the estimator fails to account for the correlation between regressors and the noise term, leading to substantial bias in the estimated coefficient of the endogenous regressor (β_2). The inefficiency share (γ) and variance components (σ_u, σ_v) also become distorted, indicating that the naive estimator is unreliable under even moderate levels of endogeneity.

The control function approach (following Kutlu, 2010) effectively corrects for endogeneity across all levels of ρ. It consistently produces unbiased estimates of both slope coefficients, including the endogenous regressor β_2, while maintaining accurate inference for inefficiency-related parameters. Its relatively low RMSE and stable convergence behavior make it the most robust estimator in the simulation. Overestimation of γ and underestimation of σ_v is only evident in high-endogeneity settings. The control function estimator strikes a favorable balance between model complexity and performance in finite samples.

The GMM estimator (following Tran and Tsionas, 2013) yields consistent estimates by relying on moment conditions derived from the joint likelihood structure. In the simulation, it successfully removes the endogeneity bias in β_2, especially when ρ is high. However, it exhibits larger RMSE and inflated estimates of γ and σ_u, along with downward bias in σ_v. These discrepancies are evident in the simulated sample size, suggesting that the method may be sensitive to finite-sample variability and the use of equally weighted moments. Despite its theoretical appeal, the GMM estimator's practical performance depends heavily on such implementation details.

The copula-based estimator (following Tran and Tsionas, 2015) addresses endogeneity by modeling the joint distribution between the endogenous regressor and the composed error, avoiding the need for external instruments. In our simulations, however, it consistently fails to correct the bias in the structural coefficient β_2, and the estimated dependence parameter ρ often deviates substantially from its true value,

[5] Because RMSE is in the same unit as the estimated parameter, it provides a more interpretable measure of estimation accuracy than MSE, which is expressed in squared units.

[6] Kutlu (2010) assumes that inefficiency is time-decaying. For simplicity of implementation, we assume that inefficiency does not decay over time.

Table 2
Monte Carlo simulation results for representative SFA models addressing endogeneity[a].

$\rho = 0$	Naive		Control function		GMM		Copula	
Parameter	bias	RMSE	bias	RMSE	bias	RMSE	bias	RMSE
β_1	−0.0018	0.0252	−0.0018	0.0252	−0.0022	0.0283	−0.0017	0.0252
β_2	−0.0006	0.0171	−0.0005	0.0247	0.0003	0.0307	0.0019	0.0811
δ			−0.0003	0.0360	−0.0013	0.0433		
γ	−0.0012	0.0459	−0.0010	0.0459	0.2137	0.2147		
σ_u	−0.0010	0.0793	−0.0009	0.0793	0.3017	0.3057	0.0122	0.0844
σ_v	−0.0014	0.0302	−0.0017	0.0302	−0.1767	0.1787	0.5389	0.5405
ρ							−0.0044	0.1349
$\rho = 0.4$								
β_1	0.0007	0.0244	0.0010	0.0233	0.0010	0.0250	0.0007	0.0244
β_2	0.2000	0.2007	0.0006	0.0255	0.0006	0.0290	0.2031	0.2173
δ			−0.0011	0.0353	−0.0010	0.0403		
γ	0.0201	0.0472	0.0426	0.0608	0.2323	0.2332		
σ_u	−0.0022	0.0728	−0.0022	0.0732	0.2447	0.2496	0.0128	0.0783
σ_v	−0.0438	0.0531	−0.0864	0.0915	−0.2485	0.2498	0.4815	0.4833
ρ							−0.4055	0.4266
$\rho = 0.8$								
β_1	0.0000	0.0227	0.0005	0.0185	0.0004	0.0192	0.0000	0.0226
β_2	0.4003	0.4006	0.0012	0.0260	0.0015	0.0269	0.4029	0.4079
δ			−0.0013	0.0319	−0.0015	0.0332		
γ	0.0937	0.1039	0.2328	0.2361	0.3252	0.3259		
σ_u	−0.0023	0.0760	−0.0031	0.0755	0.0837	0.1046	0.0169	0.0839
σ_v	−0.1779	0.1805	−0.4018	0.4029	−0.5034	0.5039	0.3052	0.3082
ρ							−0.8051	0.8145

$N = 100, T = 20, reps = 1000$; true values: $\beta_1 = \beta_2 = 0.5, \delta = 1, \gamma = 0.5, \sigma_u = \sigma_v = 1$.

[a] The Copula estimator does not include a control function term and therefore does not estimate δ. Instead, endogeneity is addressed by directly modeling the dependence between the endogenous regressor and the error term through the copula parameter ρ, which captures their joint distributional structure.
The inefficiency share $\gamma = \sigma_u^2/(\sigma_u^2+\sigma_v^2)$ is not reported for the Copula estimator, as its estimation is not based on the composed error structure typically used in standard stochastic frontier models. While σ_u and σ_v are estimated, their ratio does not carry the same interpretive meaning under the copula framework.

especially under moderate to strong endogeneity. The estimator also tends to overestimate the variance of the noise component σ_v, leading to inflated inefficiency ratios. Nonetheless, the bias and RMSE of σ_v decrease as endogeneity level increases. That said, it performs relatively well in estimating the inefficiency variance σ_u, with smaller bias and RMSE compared to GMM. This suggests that while the copula approach may struggle to adjust structural bias in small samples, it retains some accuracy in characterizing the inefficiency component of the model.

While the GMM and copula estimators show higher bias and RMSE in our simulations, these results should be interpreted in light of necessary implementation simplifications made to ensure numerical feasibility and comparability across methods.[7] For example, the GMM estimator is implemented using equally weighted moments and numerical derivatives, while the copula estimator adopts a two-stage procedure with a Gaussian copula and parametric margins. These choices may limit finite-sample accuracy but do not affect the identification structure or theoretical soundness of the estimators. As discussed above, GMM remains a robust moment-based approach when instruments are available, and the copula method offers a flexible solution in the absence of valid instruments, particularly in applications with nonlinear or unobserved dependence.

4.2. Endogeneity between the regressor, the inefficiency term and the noise term

4.2.1. Data-generating process and implementation

The *xtsfkk* estimator allows to assume correlation between the regressor and the inefficiency term, therefore, its Monte Carlo simulation has to adopt a different DGP, which basically follows Karakaplan (2022), with slight adaptations. To ensure that the simulations well capture the features of the estimator, it is realized on Stata using the *xtsfkk* command. A panel with $N = 100$ units and $T = 20$ time periods is generated as below:

$$\mathbf{y} = \beta_{c1} + \beta_{x1}\mathbf{x}_1 + \beta_{z1}\mathbf{z}_1 + \mathbf{u} + \mathbf{v},$$

$$\sigma_u^2 = \exp(\beta_{c2} + \beta_{x2}\mathbf{x}_2 + \beta_{z2}\mathbf{z}_2),$$

$$\mathbf{u}^* \sim N^+(0,1),$$

$$\mathbf{u} = \sigma_u \mathbf{u}^*,$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are exogenous variables, and $\mathbf{z}_1$ and $\mathbf{z}_2$ are endogenous variables. The true parameters are set at $\beta_{c1} = \beta_{x1} = \beta_{z1} = 0.5$, $\beta_{c2} = \beta_{x2} = \beta_{z2} = 0.25$, and all variables are generated randomly from the normal distribution with a mean of 0 and a standard deviation of 1. The endogeneity of $\mathbf{z}_1$ and $\mathbf{z}_2$ are independently and randomly generated from the normal distribution with a mean of $\mathbf{v} \times \rho$ and a standard deviation of $1-\rho$, where the degree of endogeneity increases with the ρ parameter. The IVs $\mathbf{iv}_1$ and $\mathbf{iv}_2$ are also independently and randomly generated from the normal distribution with a mean of $\mathbf{z}_1 * \delta$ and $\mathbf{z}_2 * \delta$, respectively, and a standard deviation of $1-\delta$, where the strength of IVs increases with the δ parameter. The strength of the IVs are set at levels of $\delta \in \{0.6, 0.9\}$. The Monte Carlo simulations are run for $\rho \in \{0, 0.4, 0.8\}$ with 1000 repetitions each. Results of the loops in which the endogenous models converge are used in calculating the results.

4.2.2. Simulation results and discussion

The results of Monte Carlo simulations for the *xtsfkk* estimator, including the EX — exogenous model (which ignores potential endogeneity, equivalent to the naive model in the previous subsection) and the EN — endogenous model (which applies the *xtsfkk* estimator) at three endogeneity levels and two levels of IV strength are summarized in Tables 3 and 4.

[7] Reproducing the full numerical behavior of the original estimators may require implementation in compiled languages such as C or Fortran.

Table 3
Monte Carlo simulation results for the *xtsfkk* estimator at $\delta = 0.6$.

Endogeneity	$\rho = 0$ $(attempts = 1591)$		$\rho = 0.4$ $(attempts = 1058)$		$\rho = 0.8$ $(attempts = 1060)$	
Model	EX	EN	EX	EN	EX	EN
bias β_{c1}	1.0744	1.0646	1.4150	1.1885	1.0464	1.0107
RMSE β_{c1}	1.0753	1.0657	1.4162	1.1894	1.0467	1.0111
bias β_{x1}	0.0001	0.0000	0.0003	0.0003	−0.0001	−0.0002
RMSE β_{x1}	0.0288	0.0288	0.0244	0.0248	0.0169	0.0169
bias β_{z1}	−0.0100	−0.0105	0.6725	0.7955	1.2159	1.3391
RMSE β_{z1}	0.0263	0.0326	0.6735	0.7973	1.2163	1.3395
bias β_{c2}	−3.9796	−4.5444	−1.8794	−3.2766	−4.6507	−5.0306
RMSE β_{c2}	4.0160	9.3789	1.8935	6.3122	7.1198	5.0761
bias β_{x2}	−1.0892	−1.1584	−0.5197	−0.5906	−1.1054	−1.4723
RMSE β_{x2}	1.1179	1.3073	0.5250	1.7223	1.5883	1.4818
bias β_{z2}	−1.0876	−0.8612	−2.0329	−2.0399	−1.4381	0.1063
RMSE β_{z2}	1.1149	2.8645	2.0361	2.1585	2.1600	0.4504
bias σ_v	0.2219	0.2213	0.0371	0.0325	−0.2524	−0.2750
RMSE σ_v	0.2229	0.2223	0.0411	0.0370	0.2528	0.2753

$N = 100, T = 20, reps = 1000$.

Table 4
Monte Carlo simulation results for the *xtsfkk* estimator at $\delta = 0.9$.

Endogeneity	$\rho = 0$ $(attempts = 2075)$		$\rho = 0.4$ $(attempts = 1173)$		$\rho = 0.8$ $(attempts = 1125)$	
Model	EX	EN	EX	EN	EX	EN
bias β_{c1}	1.0741	1.0572	1.4127	1.4000	1.0460	1.0252
RMSE β_{c1}	1.0751	1.0593	1.4140	1.4016	1.0463	1.0257
bias β_{x1}	−0.0010	−0.0010	−0.0000	−0.0000	0.0001	0.0000
RMSE β_{x1}	0.0291	0.0291	0.0245	0.0245	0.0170	0.0170
bias β_{z1}	−0.0103	−0.0104	0.6733	0.6780	1.2162	1.2481
RMSE β_{z1}	0.0262	0.0265	0.6742	0.6790	1.2165	1.2485
bias β_{c2}	−3.9679	4.0060	−1.8871	−2.6711	−4.6493	−9.3465
RMSE β_{c2}	−18.8113	70.6222	1.9011	13.3899	7.1189	35.5813
bias β_{x2}	−1.0828	−0.0500	−0.5228	−0.6268	−1.1035	−1.2162
RMSE β_{x2}	1.1122	15.9387	0.5280	3.7279	1.5868	3.8164
bias β_{z2}	−1.0940	−1.0698	−2.0343	−2.1116	−1.4360	0.4124
RMSE β_{z2}	1.1224	15.7230	2.0375	3.1870	2.1587	12.1557
bias σ_v	0.2220	0.2225	0.0373	0.0372	−0.2525	−0.2575
RMSE σ_v	0.2229	0.2235	0.0413	0.0416	0.2529	0.2579

$N = 100, T = 20, reps = 1000$.

In practice, there is possibility of non-convergence for the endogenous model, depending on the generated sample. The numbers of actual attempts made until 1000 valid repetitions are reached are shown in the table. When $\rho = 0$, it is more difficult for the endogenous model to converge, since the estimator is designed to address potential endogeneity. $\rho = 0.4$ and 0.8 share similar and non-ignorable frequency of non-convergence, indicating pragmatic difficulty in applying the estimator to real data. The frequency of non-convergence of all endogeneity levels grows with the strength of IV.

On the other hand, when the endogenous model converges, in rare but non-trivial cases, extreme values are estimated for some of the coefficients, with large standard deviations and statistical non-significance. In practice, when such results are obtained from the endogenous model, prudence is required in interpreting them. In normal loops, the difference between the exogenous and endogenous models are very small and insufficient to offset the extreme values. This makes the average bias and RMSE of some coefficients larger in the endogenous model. Nevertheless, we can still compare the performance of the estimator across different endogenous levels and strengths of IV.

The estimations of β_{c1} and β_{x1} are very similar in both exogenous and endogenous models, which is expected for exogenous variables. The estimations of β_{z1} have larger bias in the endogenous model when $\delta = 0.6$, but the bias becomes smaller with $\delta = 0.9$. It implies that stronger IVs may improve the performance of the *xtsfkk* estimator on endogenous variables. For the other endogenous variable, β_{z2}, the estimator outperforms the exogenous model at high endogeneity level. When $\delta = 0.9$, the average bias is smaller than $\delta = 0.6$, but RMSE grows. Meanwhile, the estimator brings about higher volatility to other variables in the inefficiency term. Considering the existence of occasional extreme values, the results of the normal estimations should be more precise. When endogeneity is at moderate level ($\rho = 0.4$), the estimations of σ_v by both models are quite good, while there is larger bias at both low ($\rho = 0$) and high ($\rho = 0.8$) endogeneity level.

Compared to the four representative estimators evaluated earlier, the *xtsfkk* estimator targets a different source of endogeneity — namely, the correlation among regressors, the inefficiency term and the noise component. This distinction is reflected in its separate data-generating process and simulation design. In terms of performance, the *xtsfkk* estimator outperforms exogenous estimator when endogeneity is present, especially in recovering coefficients of endogenous regressors. Its advantage becomes more apparent as endogeneity intensifies and instrumental variables become stronger. Nonetheless, the estimator exhibits notable sensitivity to convergence and volatility in finite samples, with occasional extreme estimates that inflate overall bias and RMSE. This trade-off highlights *xtsfkk*'s strength in addressing inefficiency-related endogeneity, while also underscoring the importance of specification choices and instrument selection.

When compared to the other estimators, *xtsfkk* offers a flexible and theoretically grounded framework suitable for panel data, with the added benefit of integrated treatment of endogeneity and inefficiency structure. However, in practice, it requires careful tuning and sometimes simplified functional forms to ensure numerical feasibility. The choice among these methods should take into account the assumed structure of endogeneity and the empirical context at hand.

Overall, the simulations show that each estimator has its own advantages and drawbacks. The naive model is clearly biased when regressors are endogenous. The control function estimator provides a straightforward correction and yields reasonably reliable results. The GMM estimator is theoretically appealing due to its general moment-based formulation, but in practice its performance advantage is limited. The copula approach offers a flexible alternative when external instruments are unavailable, with satisfactory results in finite samples. Finally, the *xtsfkk* estimator addresses endogeneity in both frontier and inefficiency terms, though convergence issues may arise. Taken together, the findings underscore that each method involves trade-offs between bias control, efficiency, and feasibility.

5. Empirical application

5.1. Empirical model

In this section, we apply the representative estimators addressing endogeneity SFA to real data of firms in the Portuguese electricity sector — specifically, the subsector of electricity generation from thermal sources. The main purpose of this section is to illustrate the application of these estimators to real data where endogeneity may exist in estimating Stochastic Frontier models, and to compare their performance in such application.

Following Kumbhakar, Wang and Horncastle (2015), the cost minimization problem for producer i under an input-oriented technical efficiency specification is

$$\min \mathbf{w}'\mathbf{x} \text{ s.t. } y = f(\mathbf{x}e^{-\eta}), \tag{29}$$

$$\text{F.O.C.: } \frac{f_j(\mathbf{x}e^{-\eta})}{f_1(\mathbf{x}e^{-\eta})} = \frac{w_j}{w_1}, j = 2, \ldots, J, \tag{30}$$

where $\mathbf{x}$ and $\mathbf{w}$ are vectors of inputs and their prices, and $\eta \geq 0$ is the input-oriented technical inefficiency that measures the percentage by which all the inputs are overused in producing output y. The cost function can therefore be defined as

$$C^*(\mathbf{w}, y) = \sum_j \mathbf{w}_j x_j e^{-\eta}, \tag{31}$$

which is the frontier cost function that gives the minimum cost given input prices $\mathbf{w}$ and the observed output level y. On the other hand, the actual cost can be written as

$$C^a = \sum_j w_j x_j = C^*(\mathbf{w}, y)\exp(\eta), \tag{32}$$

and therefore, we have

$$\ln C^a = \ln C^*(\mathbf{w}, y) + \eta. \tag{33}$$

The relationship implies that log actual cost is increased by η, i.e. all the inputs are overused by η. The efficiency index of a producer is then

$$\exp(-\eta) = \frac{C^*}{C^a}.$$

In implementation, we specifically assume that the cost function takes a translog form:

$$\begin{aligned}
\ln C^a_{it} &= \ln C^*(\mathbf{w}_{it}, y_{it}) + v_{it} + \eta_{it} \\
&= \beta_0 + \sum_j \beta_j \ln w_{jit} + \sum_j \beta_{tj} t \ln w_{jit} + \beta_y \ln y_{it} + \beta_{ty} t \ln y_{it} \\
&+ \frac{1}{2}\sum_j \sum_k \beta_{jk} \ln w_{jit} \ln w_{kit} + \frac{1}{2}\beta_{yy}(\ln y_{it})^2 + \sum_j \beta_{jy} \ln w_{jit} \ln y_{it} \\
&+ \sum_j \beta_{tj} t \ln w_{jit} + \beta_{ty} t \ln y_{it} + \beta_t t + \beta_{tt} t^2 + v_{it} \\
&+ u_{it}; v_{it} \sim i.i.d. N(0, \sigma_v^2),
\end{aligned} \tag{34}$$

where i, t denotes observation of firm i at period t, β represents unknown parameters to be estimated, v_{it} is the normally distributed error term and u_{it} is the inefficiency term that follows specific assumptions according to the estimator adopted. Some theoretical assumptions are necessary to facilitate the transformation of the cost function. Following Kumbhakar et al. (2015), $\beta_{jk} = \beta_{kj}$ is required by symmetry. The cost function is homogeneous of degree one in the input prices, which imposes the following parameter restrictions:

$$\sum_j \beta_j = 1, \sum_j \beta_{jk} = 0 \forall k, \sum_j \beta_{jy} = 0, \sum_j \beta_{tj} = 0. \tag{35}$$

Once these constraints are substituted into the model, the homogeneity conditions are automatically satisfied. This procedure amounts to using one of the input prices to normalize cost and other input prices. With K and L representing capital and labor as two inputs, after substitution and manipulation, we normalize the cost function using w_{Kit} as the normalizing price, obtaining

$$\begin{aligned}
\ln(\frac{C^a_{it}}{w_{Kit}}) &= \beta_0 + \beta_y \ln y_{it} + \beta_{ty} t \ln y_{it} + \beta_L \ln(\frac{w_{Lit}}{w_{Kit}}) + \beta_{tL} t \ln(\frac{w_{Lit}}{w_{Kit}}) \\
&+ \frac{1}{2}\beta_{yy}(\ln y_{it})^2 \\
&+ \frac{1}{2}\beta_{LL}(\frac{w_{Lit}}{w_{Kit}})^2 + \beta_{Ly} \ln(\frac{w_{Lit}}{w_{Kit}}) \ln y_{it} + \beta_t t + \frac{1}{2}\beta_{tt} t^2 + v_{it} + u_{it}.
\end{aligned} \tag{36}$$

The equation above is estimated with the 5 estimators described in Section 3: the naive estimator, the control function estimator, the GMM estimator, the copula-based estimator and the *xtsfkk* estimator. Necessary adjustments are made according to the cost function.

5.2. Data

The data employed in the empirical application is part of the BPLIM database[8] of the Bank of Portugal (Banco de Portugal). Firms are identified by anonymized tax/bank identification numbers and the data can only be accessed on BPLIM's remote servers. The data used in this study comes from the Central Balance Sheet, mostly based on information reported through Informação Empresarial Simplificada (IES, Simplified Corporate Information) and contains annual data. As the database is not exclusively dedicated to the electricity sector, it contains general information on firms' financial status.

Depending on data availability, we choose a set of variables to estimate Stochastic Frontier cost functions with annual panel data from 2006 to 2021 for firms in thermal power subsector of Portugal. The following variables are used in our study:

Frontier equation:

- y - measured by non-financial revenue;
- w_K - calculated by interest expenses divided by obtained funding as a proxy for the price of capital;
- w_L - measured by average hourly wage, calculated by salaries paid to employees divided by total hours worked;
- C^a - calculated by the sum of financial expenses, salary expenses and expenses on goods and materials.

Inefficiency variables (when applicable):

- Age (AGE): the age of the firm; the impact of firm age on technical inefficiency is studied by Lai and Kumbhakar (2018);
- Financial investment (FIV): measured by the natural logarithm financial investment; firms' involvement in financial activities may affect their efficiency (Hou et al., 2021, 2024).

As in the *xtsfkk* estimator, endogenous variables can be defined in both frontier and inefficiency equations, we assume one endogenous variable in each of them. For firms operating in electricity generation, it is reasonable to assume that the output is determined by demand and is thus exogenous (Liu et al., 2019). Therefore, $\ln W = \ln(\frac{w_L}{w_K})$ is assumed as the endogenous variable in the frontier equation. Among the

[8] Website: https://bplim.bportugal.pt/.

Table 5
Descriptive statistics of data of the thermal power subsector of Portugal.
Source: Descriptive statistics based on data from the BPLIM database.

Variable	Unit	Obs.	Mean	Std. Dev.	Missing %
y	Euro	2468	$1.96e+07$	$1.23e+08$	0%
w_K	Ratio	1006	.1561628	1.170435	59.23%
w_L	Euro/hour	898	13.13717	13.22797	63.61%
C^a	Euro	2468	2 201 233	$1.90e+07$	0%
Age	Year	2468	12.17018	9.814504	0%
FIV	Euro	666	5 285 540	$3.51e+07$	73.01%
CD	Ratio	829	4 661 586	$3.32e+07$	66.41%
$AVHRS$	Hour	896	1716.052	446.9426	63.69%

Note: Minimum and maximum values omitted as requested by confidentiality terms of the BPLIM database.

inefficiency variables, a firm's age is apparently exogenous; its financial investment (FIV), on the other hand, can be related to unobserved operational features of each firm and thus assumed as endogenous. In the control-function estimator, the GMM estimator and the copula-based estimator, only $\ln W$ is assumed endogenous since the specifications do not allow explanatory variables for the inefficiency term.

Instrumental variables[9] (when applicable):

- Capital deepening (CD): measured by the natural logarithm of the ratio of capital (fixed tangible asset) to labor (number of employees); the ratio of capital to labor of a firm may reflect a firm's incentive for financial investment (FIV); it may also indicate willingness to pay higher wage to its employees, thus be related to $\ln W$;
- Average hours worked (AVHRS): measured by the natural logarithm of average hours worked per paid employee; working hours can reflect a firm's managerial structure and intuitively, it can be related to average wage (therefore $\ln W$).

Both IVs are assumed for the control-function estimator, the GMM estimator and the *xtsfkk* estimator. The copula-based estimator does not require IVs.

Descriptive statistics of the data of each subsector in our study are presented in Table 5. The descriptive statistics are based on the original values of the variables. As some of the observations take the original values of 0, missing values are generated when transformed into natural logarithms. This makes the panel less balanced and reduces the effective observations in the estimation. The percentage of missing values of each variable is also reported in the table.

Additionally, the correlation matrix among the variables used in the estimation is reported in Table 6. The matrix is calculated with the values actually used in the estimation, i.e., after transformations such as taking natural logarithms.

The correlation matrix indicate that CD is more strongly correlated with the endogenous wage variable than $AVHRS$. Such correlations offer preliminary intuition about instrument relevance, but the strength of these IVs requires formal test discussed in the next subsection.

5.3. Results and discussion

The empirical model described in Section 5.1 is estimated with the estimators mentioned in Section 3. For each estimation procedure, several diagnostics are embedded. First, an endogeneity test is done according to Section 3.5 to detect the presence of potential endogeneity. Second, weak instruments are assessed both in the R implementation and in the *xtsfkk* framework. In R, a first-stage regression of the

[9] The effects of the above inefficiency and/or instrumental variables may imply firm-specific production/cost features and thus be confounded with the residual forming the cost inefficiency. While there are models dedicated to treating such effect, the analysis of this section focus on the endogeneity issue.

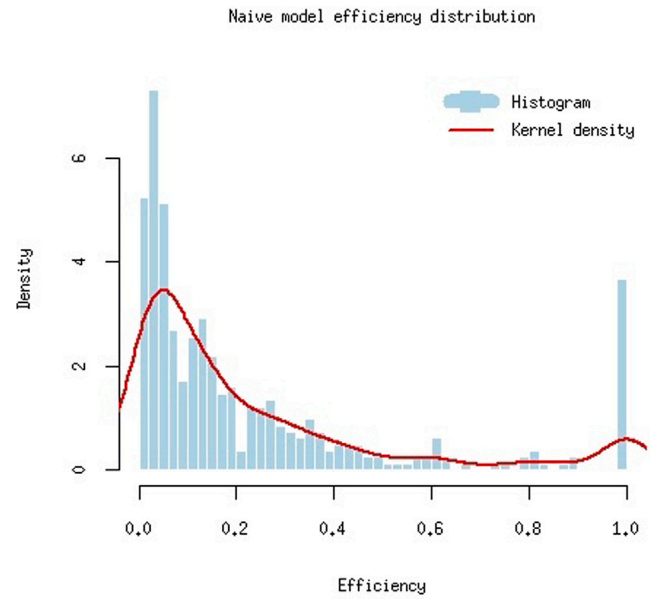

Fig. 1. Distribution of cost efficiency in the Portuguese thermal power subsector, Naive estimator.

endogenous variable on the instruments and exogenous covariates is estimated, and the joint significance of the instruments is tested. In *xtsfkk*, the test command provides an analogous check, yielding a chi-squared statistic for the null that a given instrument has no explanatory power across all endogenous regressors. In both cases, rejection of the null indicates instrument relevance. Third, in the R framework, since the number of instruments exceeds the number of endogenous regressors, an overidentification test is conducted. Specifically, a Sargan nR^2 statistic, which tests the null hypothesis that the instruments are valid. In contrast, the *xtsfkk* specification with two instruments for two endogenous regressors is exactly identified, and thus no overidentification test is available.

Similar to the Monte Carlo simulations, the *xtsfkk* estimator is applied using Stata. The other estimators are implemented using R, and the results are presented in Table 7.

The results presented in Table 7 are estimated assuming CD and AVHRS as IVs; estimated results, including test results, are very similar if only one of them is assumed as IV (although the model is supposed to be correctly identified with one endogenous variable and one IV). Endogeneity tests cannot reject the null hypothesis that there is no endogeneity in the model, which is possibly because of the high standard error in the data, and thus, in the test results. Weak IV tests show that the IVs adopted in the model is strong enough. When two IVs are assumed, the model is likely to be over-identified; nonetheless, there is marginal difference in the results compared with the those when one IV is assumed.

In the naive model, without considering endogeneity, coefficients on $\ln y$ and $(\ln y)^2$ are statistically significant. In the control-function model, the coefficient on $(\ln y)^2$ loses statistical significance while that on the endogenous variable $\ln W$ gains it. As demonstrated by the Monte Carlo simulations, the naive estimator generates large biases with endogenous variables (when endogeneity is present). By accounting for endogeneity, such biases are corrected, leading to changes in statistical significances of the coefficients. This is in line with theoretical expectations. In the GMM model, the sign on $\ln W$ is reversed, which is a typical symptom of endogeneity. Although collinearity among regressors can amplify this effect, the primary source is the correlation between explanatory variables and the error term.

The GMM estimator delivers relatively large standard errors and weaker significance, indicating limited precision in finite samples. This

Table 6
Correlation matrix among the variables used in the empirical estimation.
Source: Reported by Stata based on data from the BPLIM database.

	$\ln C$	$\ln y$	$\ln W$	AGE	FIV	CD	$AVHRS$
$\ln C$	1.0000						
$\ln y$	0.4696***	1.0000					
$\ln W$	0.6970***	0.0892*	1.0000				
AGE	0.2725***	0.1676***	−0.0174	1.0000			
FIV	0.5365***	0.3596***	0.1188*	0.1870***	1.0000		
CD	0.2216***	0.2752***	0.1597***	0.0991***	0.1297***	1.0000	
$AVHRS$	0.1442***	0.1334***	−0.0769	0.0551*	0.0024	−0.0670*	1.0000

Note: $\ln C = \ln(C^a/w_K)$; $\ln W = w_L/w_K$;.
*/**/*** stands for statistical significance at 10%/5%/1% level.

Table 7
Estimated results of the naive estimator, the cost function estimator, the GMM estimator and the copula-based estimator.

Variable	Coefficient			
	Naive	Control function[a]	GMM[a]	Copula
Frontier				
$\ln y$	−1.7077***(0.1501)	−1.5302***(0.5037)	−0.7010**(0.3491)	−1.5249***(0.1514)
$\ln W$	−0.1067(0.7611)	−0.4515***(0.0213)	0.0039(1.3308)	−0.3466(0.8274)
$(\ln y)^2$	0.0734***(0.0054)	0.0708(1.7371)	0.1746**(0.0779)	0.0697***(0.0055)
$(\ln W)^2$	−0.0914(0.0775)	−0.1085(0.3631)	0.4175(0.9748)	−0.1636**(0.0810)
$\ln y \ln W$	0.0588(0.0504)	0.0431(0.2292)	−0.2531(0.3517)	0.0372(0.0520)
$t \ln y$	0.0054(0.0099)	−0.0021(0.0292)	0.4431***(0.0569)	0.0001(0.0097)
$t \ln W$	−0.0223(0.0325)	−0.0242(0.0922)	−0.0807(0.3841)	−0.0115(0.0331)
t	−0.0943(0.1565)	0.0534(0.5075)	0.3151(0.6981)	−0.0002(0.1545)
t^2	0.0038(0.0052)	0.0022(0.0106)	−0.2261***(0.0261)	0.0014(0.0052)
Intercept	21.8979***(2.0487)	21.2599***(2.7972)	21.5928***(2.1590)	23.7919***(2.3417)
$\ln \sigma_u$	0.5517***(0.1720)	0.5901(0.4962)	0.1158(0.6417)	1.0607***(0.0573)
$\ln \sigma_v$	0.5023***(0.0716)	0.4825***(0.1748)	0.1190(0.5437)	0.4636***(0.1597)
δ		0.7143(2.6260)	0.2527(1.5247)	
ρ				−0.3194(1.1949)
Efficiency				
Mean	0.2309	0.4745	0.6893	0.2828
Median	0.1235	0.4807	0.9999	0.2836
Endogeneity test				
Wald		0.07400148	0.02747196	0.8725226
p-value		0.7855973	0.8683563	0.3502579
Weak IV test: $F = 10.577, Pr(>F)$=3.342$e-05$ (Reject H_0 : $CD = AVHRS = 0$)				
Over-identification test: Sargan $nR^2 = 1.1716, p = 0.2791$ (Cannot reject H_0; model over identified.)				

Note: $N.obs.$=410; $\ln W = \ln(w_L/w_K)$; */**/*** stands for statistical significance at 10%/5%/1% level.
[a] Standard errors calculated by bootstrap with 100 repetitions.

outcome is consistent with the Monte Carlo simulations, where GMM was shown to be robust in principle but less efficient with the simulated sample size. By contrast, the copula estimator yields results similar to those of the control-function model, further confirming its ability to account for endogeneity by capturing the dependence structure between regressors and the error component. Taken together, these empirical results align with the simulation evidence: while the naive estimator suffers from bias, both the control-function and copula approaches correct it, and GMM remains less precise despite its general robustness.

Estimated results on $\ln \sigma_u$ and $\ln \sigma_v$ demonstrate apparent differences across these approaches. Similar to simulation results, the GMM results diverge substantially from the other estimators, with much larger standard errors. While the estimates of $\ln \sigma_v$ are fairly consistent across the other estimators, those of $\ln \sigma_u$ differ case by case. The control-function estimator provides a coefficient similar to the naive estimator but with larger standard error; the copula estimator provides a larger coefficient with smaller standard error.

The predicted efficiency values of each estimator should be interpreted along with Figs. 1–4, which demonstrate the histogram and kernel density of predicted efficiency levels for each of the estimators.

In the naive model, the mean efficiency score is relatively low, with an even lower median, and the overall distribution is left-skewed. In addition, a large number of observations are concentrated at full efficiency (100%).[10]

In the control-function model, the mean efficiency is higher and close to the median, but the distribution is irregular, exhibiting bimodality with two distinct modes. In fact, efficiency values predicted by the naive model can also be considered as bimodal, but with a more dispersed pattern; by reducing the mass at full efficiency, the control-function model brings the two modes closer together and thereby largely corrects the bias.

In the GMM model, the mean efficiency is about 0.6893, higher than in the previous two models, while the median reaches 0.9999. This apparent anomaly is explained by the distribution in Fig. 3: a small number of observations lie in the middle-to-low range, whereas a large mass is truncated at full efficiency (unity). This outcome contrasts with the large bias observed in the estimate of $\ln \sigma_u$, further highlighting the instability of the GMM estimator.

In the copula model, the mean and median efficiency values are very close, lying between those of the naive and control-function models.

[10] Efficiency values exceeding unity occur when the conditional expectation of inefficiency, $E(u_i|\epsilon_i)$, becomes negative due to estimation noise, which after transformation yields $\exp(-E(u_i|\epsilon_i)) > 1$; these values are truncated at one.

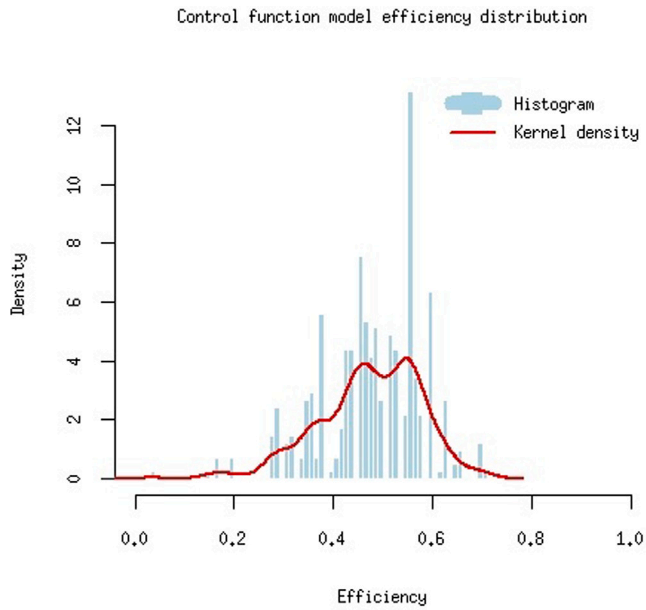

Fig. 2. Distribution of cost efficiency in the Portuguese thermal power subsector, control-function estimator.

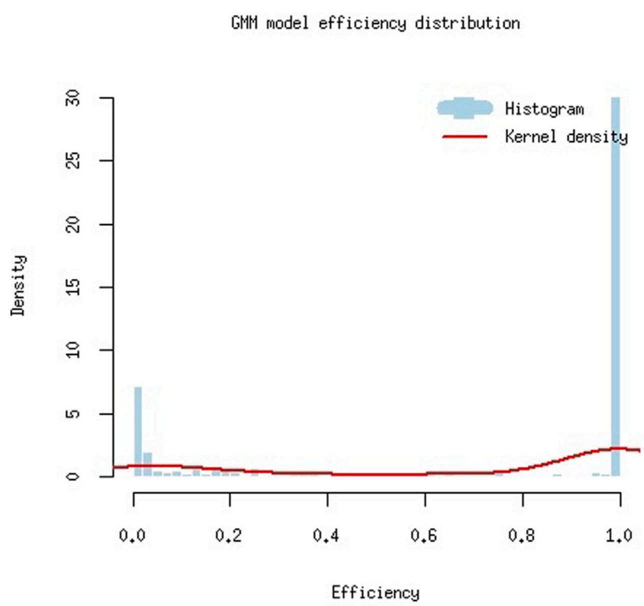

Fig. 3. Distribution of cost efficiency in the Portuguese thermal power subsector, GMM estimator.

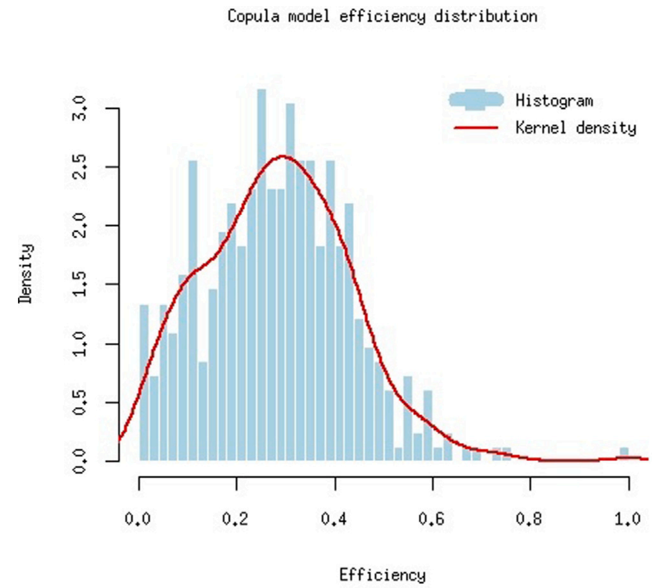

Fig. 4. Distribution of cost efficiency in the Portuguese thermal power subsector, copula-based estimator.

Table 8
Estimated results of the *xtsfkk* estimator.

Variable	Coefficient	
	Exogenous (EX) model	Endogenous (EN) model
Frontier		
$\ln y$	0.2017(0.2580)	-1.5545^{***}(.4642)
$\ln W$	0.3974^{**}(.1791)	-4.3991^{***}(1.6162)
$(\ln y)^2$	0.0080(0.0089)	0.0120(0.0145)
$(\ln W)^2$	0.0145^{***}(0.0035)	0.0093(0.0090)
$\ln y \ln W$	0.0092(0.0132)	0.3294^{***}(0.1002)
$t \ln y$	-0.0426^{***}(0.0052)	-0.0624^{***}(0.0141)
$t \ln W$	0.0062(0.0060)	0.0115(0.0139)
t	0.6347^{***}(0.0863)	0.8993^{***}(0.2109)
t^2	-0.0076^{***}(0.0018)	-0.0084^{*}(0.0043)
Intercept	5.8306^{***}(1.9172)	31.2422^{***}(6.7084)
Inefficiency term		
AGE	0.0894^{***}(0.0112)	0.0918^{***}(0.0110)
FIV	-0.0128(0.0186)	-0.1705^{***}(0.0295)
Intercept	0.7093^{***}(0.3234)	2.5019^{***}(0.5043)
Efficiency		
Mean	0.1988	0.2135
Median	0.0993	0.1260
Endogeneity test: $\chi^2(2) = 23.06(prob > \chi^2 = 0.0000)$		
Reject $H_0 : \eta_{\ln W} = \eta_{FIV} = 0$ at 1% level.		
Weak IV test		
CD	$\chi^2(2) = 44.57(prob > \chi^2 = 0.0000)$	
$AVHRS$	$\chi^2(2) = 6.53(prob > \chi^2 = 0.0381)$	

Note: N.obs.=234; $\ln W = \ln(w_L/w_K)$;
*/**/*** stands for statistical significance at 10%/5%/1% level.

The distribution is approximately normal, and the bimodality observed in the other models is much less pronounced.

The comparison of efficiency distributions reveals that both the naive and control-function models exhibit bimodality, though the latter reduces the mass at full efficiency and brings the two modes closer together, largely correcting the bias. The GMM estimator, by contrast, produces an anomalous pattern with most observations truncated at unity, consistent with its unstable variance estimates. The copula model stands out by yielding a smoother, near-normal distribution with close mean and median values, suggesting a more balanced characterization of efficiency.

The Stata command *xtsfkk* reports results for the exogenous model (where endogeneity is ignored) and the endogenous model (where the estimator is used to correct for endogeneity). The results of both are presented in Table 8.

∂ Endogeneity test unambiguously confirms the presence of endogeneity in the model. Although in Table 7 the evidence for endogeneity is weaker, this likely reflects differences in model specification rather than its true absence. It is therefore reasonable to conclude that endogeneity remains a relevant concern, and that accounting for it materially changes both the coefficients and the efficiency estimates. χ^2 statistics show that both IVs are valid. In this estimation, 234 effective observations are used, fewer than the number used for Table 7. With

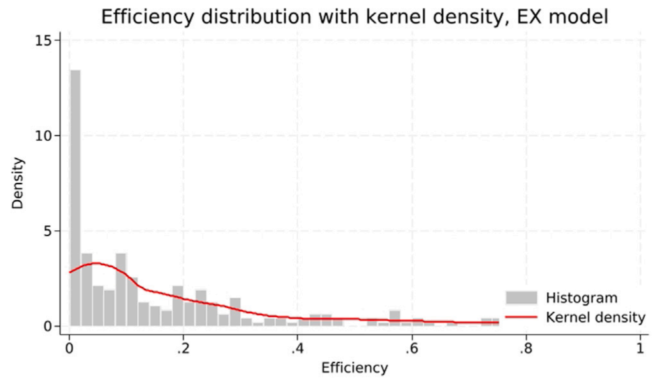

Fig. 5. Distribution of cost efficiency in the Portuguese thermal power subsector, exogenous model estimated with *xtsfkk* package.

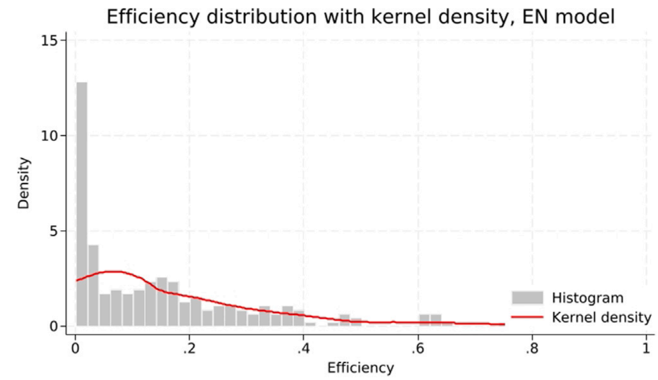

Fig. 6. Distribution of cost efficiency in the Portuguese thermal power subsector, endogenous model estimated with *xtsfkk* package.

missing values for some variables in the data, more complicated specifications lead to fewer effective observations, which may undermine the accuracy of the results. This is another factor to be taken into account in practical application of estimating approaches.

In Table 8, the endogenous specification produces notable changes relative to the exogenous one. The coefficients on $\ln y$ and $\ln W$ switch signs, a pattern also seen in the GMM results of Table 7, underscoring how accounting for endogeneity alters the estimated relationships. For the inefficiency determinants, FIV is statistically insignificant in the exogenous model but becomes significant once endogeneity is controlled for. The estimated results on inefficiency explanatory variables AGE and FIV[11] are consistent with previous research (Hou et al., 2024). A firm with longer history may suffer from inertia in its managerial structure and face more difficulty in cost optimization. The negative association between FIV and inefficiency suggests that firms with greater financial investment tend to operate more efficiently. A possible explanation is that stronger financial capacity facilitates access to better technologies, management practices, or scale economies, thereby reducing inefficiency.

The abovementioned differences show that the *xtsfkk* estimator can correct biases caused by endogeneity in both frontier and inefficiency equations. Efficiency scores also differ markedly: the endogenous model yields higher mean and median values, indicating that efficiency is likely underestimated when endogeneity is ignored. The distribution of efficiency values of the exogenous and endogenous models are illustrated in Figs. 5 and 6.

The distribution of efficiency values predicted by the exogenous and endogenous models are similar; the efficiency distribution under the EN specification appears slightly right-skewed and with lower kurtosis compared to the EX model. The efficiency distribution of these models more closely resembles that of the naive model (rather than that of the cost-function or copula model). This pattern may reflect incomplete correction on the predicted efficiency scores under the influence of inefficiency determinants.

The representative models examined in this paper each offer distinct advantages in addressing endogeneity. The control-function approach is relatively simple to implement and delivers reliable results. The GMM framework contributes more on the theoretical side than in empirical performance. The copula model, by not requiring instruments, provides reasonably robust estimates and may serve as a flexible alternative in some contexts. The *xtsfkk* estimator extends further by incorporating inefficiency determinants, enhancing practical applicability; its Stata package greatly facilitates its usage in applied studies. Nevertheless, the inclusion of too many variables can increase convergence difficulties and destabilize efficiency estimates, which calls for cautious application in practice. The replication package includes an R script which implements estimators applied in this empirical application (except the *xtsfkk* estimator), which can be used as a reference for practical studies with proper adaptations. Issues such as presence of endogeneity, instrument strength, and over-identification are highly sensitive to model specification, and test outcomes can vary substantially depending on how the model is defined. Hence, careful tuning is essential in empirical applications.

From a policy perspective, the empirical application illustrates how different estimators can alter the interpretation of efficiency patterns in the thermal power subsector. Variations in average efficiency levels, the significance of firm characteristics such as AGE and FIV, and the ranking of firms by efficiency scores are not merely econometric details, but indicators that regulators may use when setting revenue caps, designing incentive schemes, or prioritizing investment support. The finding that financial investment (FIV) reduces inefficiency, for instance, points to the importance of facilitating firms' access to capital in order to promote technological upgrading and cost optimization. Conversely, the sensitivity of results to estimator choice cautions against relying on a single benchmark when formulating regulation. Thus, methodological rigor in addressing endogeneity directly translates into more reliable policy-relevant indicators for the governance of the energy sector.

6. Conclusion

Stochastic Frontier Analysis is widely applied in the energy sector, as efficiency scores derived from frontier models often serve as references for regulatory policies aimed at stimulating competition and improving performance. Yet, the presence of endogeneity between input variables and the inefficiency component can seriously compromise the accuracy of such evaluations. Because of the specific structure of frontier models, dedicated estimators are needed to address this issue. A range of approaches have recently been proposed, including IV-based maximum likelihood estimators, GMM, and copula models, each offering a different route to mitigate the bias induced by endogeneity.

This study evaluates the performance of alternative estimators through Monte Carlo simulations and an empirical application to the Portuguese thermal power subsector from 2006 to 2021. The results indicate that neglecting potential endogeneity can lead to substantial differences in both frontier coefficients and efficiency scores. Models that explicitly account for endogeneity generally deliver more robust and interpretable outcomes, while exogenous specifications risk understating inefficiency and mischaracterizing its determinants. Overall, addressing endogeneity is crucial to ensure that such policy measures are based on reliable efficiency benchmarks.

[11] Since the definition of the inefficiency term is quite complex, as defined by Eq. (22), it is difficult to intuitively interpret the magnitude of the effects of these variables on inefficiency. Nevertheless, the sign and statistical significance is of interest for policy consideration.

Turning to the empirical findings, firm age is generally associated with higher inefficiency, reflecting inertia in managerial structures. Financial investment reduces inefficiency, pointing to the importance of access to capital in facilitating technological upgrading and cost optimization.

Several limitations of this study should be acknowledged. First, the database employed is not specifically designed for the energy sector, and some relevant variables are either imperfectly measured, unavailable, or subject to missing values. Second, these data limitations constrain the set of inefficiency determinants that can be incorporated, which in turn narrows the scope of policy insights that can be drawn from the empirical results. Third, our approach addresses endogeneity at the econometric level rather than through a structural cost specification that disentangles modeling errors from price-related inefficiency, partly because the available data does not permit such modeling. Future research may overcome these shortcomings by using more specialized and higher-quality datasets. While these limitations restrict the extent of sector-specific policy contributions, they do not undermine the main aim of the paper, which is to provide a comparative assessment of alternative estimators for handling endogeneity in Stochastic Frontier models.

This analysis also offers guidance for further applications. Flexible specifications such as translog may improve realism but often raise convergence issues, making simpler forms preferable in practice. Future research could combine endogeneity-corrected estimators with richer sectoral datasets that incorporate regulatory measures, thereby linking efficiency outcomes more closely to policy. Extending the comparative assessment to other industries or frontier frameworks would further strengthen the methodological toolkit for applied efficiency analysis.

CRediT authorship contribution statement

Zheng Hou: Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Joaquim J.S. Ramalho:** Writing – review & editing, Methodology. **Catarina Roseta-Palma:** Writing – review & editing.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to optimize the programming and to polish the language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Acknowledgments

We appreciate the comments of the reviewers and editors, which greatly helped us improve the article. We thank the BPLIM team for the support and assistance with the database. Zheng Hou is financed by Portuguese public funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., in the framework of the project UID/04105/2023 and UIDB/05069/2020 (REM/UECE). Joaquim J.S. Ramalho and Catarina Roseta-Palma are supported by FCT - Fundação para a Ciência e a Tecnologia (Grant UIDB/00315/2020).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eneco.2025.108922.

References

Aigner, D.J., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. J. Econometrics 6 (1), 21–37.
Alizadeh, R., Beiragh, R.G., Soltanisehat, L., Soltanzadeh, E., Lund, P.D., 2020. Performance evaluation of complex electricity generation systems: A dynamic network-based data envelopment analysis approach. Energy Econ. 91, 104894.
Amanda, C., 2023. Spatial competition on rural bank efficiency: Evidence from Indonesia. Cogent Econ. Financ. 11 (1), 2190216.
Amsler, C., Prokhorov, A., Schmidt, P., 2016. Endogeneity in stochastic frontier models. J. Econometrics 190 (2), 280–288.
Amsler, C., Prokhorov, A., Schmidt, P., 2017. Endogenous environmental variables in stochastic frontier models. J. Econometrics 199 (2), 131–140.
Barros, C.P., 2008. Efficiency analysis of hydroelectric generating plants: A case study for Portugal. Energy Econ. 30, 59–75.
Battese, G.E., Coelli, T.J., 1992. Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. J. Prod. Anal. 3, 153–169.
Belotti, F., Ferrara, G., 2021. Imposing Monotonicity in Stochastic Frontier Models: An Iterative Nonlinear Least Squares Procedure. In: CEIS Tor Vergata Research Paper Series, vol. 17, p. 462, No. 5.
Bigerna, S., D'Errico, M.C., Polinori, P., 2019. Environmental and energy efficiency of EU electricity industry: An almost spatial two stages DEA approach. Energy J. 40 (1), 31–56.
Bigerna, S., D'Errico, M.C., Polinori, P., 2020. Heterogeneous impacts of regulatory policy stringency on the EU electricity industry: A Bayesian shrinkage dynamic analysis. Energy Policy 142, 111522.
Bigerna, S., D'Errico, M.C., Polinori, P., 2022. Sustainable power generation in europe: A panel data analysis of the effects of market and environmental regulations. Environ. Resour. Econ. 83, 445–479.
Binh, T.V., Mulatu, A., Xu, B., 2022. Corporate social responsibility & firm efficiency: Evidence from endogenous cost inefficiency stochastic frontier analysis. Appl. Econ. 54 (55), 6380–6392.
Bobde, S.M., Tanaka, M., 2020. Structural reforms and technical efficiency in the Indian electricity distribution sector. Asia-Pac. J. Oper. Res. 37 (2), 2050004.
Deng, Z., Zhao, Q., Bao, H.X.H., 2020. The impact of urbanization on farmland productivity: Implications for China's requisition–compensation balance of farmland policy. Land 9 (9), 311.
Diewart, W.E., Nakamura, A.O., 1999. Benchmarking and the measurement of best practice efficiency: An electricity generation application. Can. J. Econ. / Rev. Can. D'Economique 32 (2), 570–588, Special Issue on Service Sector Productivity and the Productivity Paradox.
Farrell, M.J., 1957. The measurement of productive efficiency. J. R. Stat. Soc. Ser. 1 120, 251–281.
Fungáčová, Z., Kerola, E., Weill, L., 2023. Does bank efficiency affect the bank lending channel in China? Emerg. Mark. Rev. 55, 100964.
Germeshausen, R., Panke, T., Wetzel, H., 2020. Firm characteristics and the ability to exercise market power: Empirical evidence from the iron ore market. Empir. Econ. 58, 2223–2247.
Griffiths, W.E., Hajargasht, G., 2016. Some models for stochastic frontiers with endogeneity. J. Econometrics 190 (2), 341–348.
Growitsch, C., Jamasb, T., Wetzel, H., 2012. Efficiency effects of observed and unobserved heterogeneity: Evidence from Norwegian electricity distribution networks. Energy Econ. 34, 542–548.
Guan, Z., Kumbhakar, S.C., Myers, R.J., Lansink, A.O., 2009. Measuring excess capital capacity in agricultural production. Am. J. Agric. Econ. 91 (3), 765–776.
Gultom, Y.M.L., 2019. Governance structures and efficiency in the U.S. electricity sector after the market restructuring and deregulation. Energy Policy 129, 1008–1019.
Haschka, R.E., Herwartz, H., 2022. Endogeneity in pharmaceutical knowledge generation: An instrument-free copula approach for Poisson frontier models. J. Econ. Manag. Strat. 31 (4), 942–960.
Hou, Z., Roseta-Palma, C., Ramalho, J., 2020. Directed technological change, energy and more: A modern story. Environ. Dev. Econ. 25 (6), 611–633.
Hou, Z., Roseta-Palma, C., Ramalho, J., 2021. Does directed technological change favor energy? Firm-level evidence from Portugal. Energy Econ. 98, 105248.
Hou, Z., Roseta-Palma, C., Ramalho, J., 2024. Can technical efficiency in the portuguese electricity sector be improved? Yes, but.... Energy Policy 190, 114146.
Jamasb, T., Pollitt, M., 2001. Benchmarking and regulation: international electricity experience. Util. Policy 9, 107–130.
Jerop, R., Owuor, G., Mshenga, P., Kimurto, P., 2020. Effects of finger millet innovations on productivity in Kenya. Cogent Food Agric. 6 (1), 1830476.
Jindal, A., Nilakantan, R., 2021. Falling efficiency levels of Indian coal-fired power plants: A slacks-based analysis. Energy Econ. 93, 105022.
Jondrow, J., Lovell, C.A.K., Materov, I.S., Schmidt, P., 1982. On the estimation of technical inefficiency in the stochastic frontier production function model. J. Econometrics 19 (2–3), 233–238.
Karakaplan, M.U., 2017. Fitting endogenous stochastic frontier models in Stata. Stata J. 17 (1), 39–55.
Karakaplan, M.U., 2022. Panel stochastic frontier models with endogeneity. Stata J. 22 (3), 643–663.

Karakaplan, M.U., Kutlu, L., 2017a. Handling endogeneity in stochastic frontier analysis. Econ. Bull. 37 (2), 889–901.
Karakaplan, M.U., Kutlu, L., 2017b. Endogeneity in panel stochastic frontier models: an application to the Japanese cotton spinning industry. Appl. Econ. 49 (59), 5935–5939.
Karakaplan, M.U., Kutlu, L., 2019. School district consolidation policies: Endogenous cost inefficiency and saving reversals. Empir. Econ. 56, 1729–1768.
Kim, C.-J., 2006. Time-varying parameter models with endogenous regressors. Econ. Lett. 91 (1), 21–26.
Kim, Y., Kim, C.-J., 2011. Dealing with endogeneity in a time-varying parameter model: Joint estimation and two-step estimation procedures. Econom. J. 14 (3), 487–497.
Kumbhakar, S.C., Amundsveen, R., Kvile, H.M., Lien, G., 2015. Scale economies, technical change and efficiency in Norwegian electricity distribution. J. Prod. Anal. 43, 295–305, 1998–2010.
Kumbhakar, S.C., Bernstein, D.H., 2019. Does xistence of inefficiency matter to a neoclassical xorcist? Some econometric issues in panel stochastic frontier models. In: Methodological Contributions To the Advancement of Productivity and Efficiency Analysis.
Kumbhakar, S.C., Lien, G., 2017. Yardstick regulation of electricity distribution—Disentangling short-run and long-run inefficiencies. Energy J. 38 (5), 17–37.
Kumbhakar, S.C., Mydland, Ø., Musau, A., Lien, G., 2020. Disentangling costs of persistent and transient technical inefficiency and input misallocation: The case of norwegian electricity distribution firms. Energy J. 41 (3), 143–160.
Kumbhakar, S.C., Tsionas, E.G., 2008. Scale and efficiency measurement using a semiparametric stochastic frontier model: Evidence from the U.S. commercial banks. Empir. Econ. 34, 585–602.
Kuosmanen, T., 2023. Control function in convex regression: Application to benchmarking of finish DSOs. In: Workshop on Productivity and Efficiency Analysis. Oslo, Norway.
Kutlu, L., 2010. Battese-Coelli estimator with endogenous regressors. Econ. Lett. 109 (2), 79–81.
Kutlu, L., 2022. Spatial stochastic frontier model with endogenous weighting matrix. Empir. Econ. 63 (4), 1947–1968.
Kutlu, L., Tran, K., Tsionas, M., 2019. A time-varying true individual effects model with endogenous regressors. J. Econometrics 211 (2), 539–559.
Lai, H., Kumbhakar, S.C., 2018. Endogeneity in panel data stochastic frontier model with determinants of persistent and transient inefficiency. Econ. Lett. 162, 5–9.
Lai, H., Kumbhakar, S.C., 2019. Technical and allocative efficiency in a panel stochastic production frontier system model. European J. Oper. Res. 278 (1), 255–265.
Lee, J.M., Howard, G., 2021. The impact of technical efficiency, innovation, and climate policy on the economic viability of renewable electricity generation. Energy Econ. 100, 105357.
Lien, G., Kumbhakar, S.C., Alem, H., 2018. Endogeneity, heterogeneity, and determinants of inefficiency in norwegian crop-producing farms. Int. J. Prod. Econ. 201, 53–61.
Liu, X.Y., Pollitt, M.G., Xie, B.C., Liu, L.Q., 2019. Does environmental heterogeneity affect the productive efficiency of grid utilities in China? Energy Econ. 83, 333–344.
Llorca, M., Baños, J., Somoza, J., Arbués, P., 2017. A stochastic frontier analysis approach for estimating energy demand and efficiency in the transport sector of Latin America and the Caribbean. Energy J. 38 (5), 153–174.
Lundin, E., 2020. Effects of privatization on price and labor efficiency: The Swedish electricity distribution sector. Energy J. 41 (2), 247–273.
Ma, C., Zhao, X., 2015. China's electricity market restructuring and technology mandates: Plant-level evidence for changing operational efficiency. Energy Econ. 47, 227–237.
Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from Cobb–Douglas production functions with composed error. Internat. Econom. Rev. 18 (2), 435–444.
Melo-Becerra, L.A., Ramírez-Giraldo, M.T., 2023. Transport infrastructure and technical efficiency in a panel of countries: accounting for endogeneity in a stochastic frontier model. SN Bus. Econ. 3, 34.
Mirza, F.M., Rizvi, S.B., Bergland, O., 2021. Service quality, technical efficiency and total factor productivity growth in Pakistan's post-reform electricity distribution companies. Util. Policy 68, 101156.
Musau, A., Kumbhakar, S.C., Mydland, Ø., Lien, G., 2021. Determinants of allocative and technical inefficiency in stochastic frontier models: An analysis of norwegian electricity distribution firms. European J. Oper. Res. 288 (3), 983–991.
Mydland, Ø., Haugom, E., Lien, G., 2018. Economies of scale in norwegian electricity distribution: A quantile regression approach. Appl. Econ. 50 (40), 4360–4372.
Nakaishi, T., Takayabu, H., Eguchi, S., 2021. Environmental efficiency analysis of China's coal-fired power plants considering heterogeneity in power generation company groups. Energy Econ. 102, 105511.
Navarro-Chávez, C.L., Delfín-Ortega, O.V., Díaz-Pulido, A., 2020. Efficiency of the electricity sector in Mexico 2008–2015: An application of the DEA network model. Int. J. Energy Sect. Manag. 14 (4), 683–706.
Ojo, T.O., Baiyegunhi, L.J.S., 2020. Impact of climate change adaptation strategies on rice productivity in South-west, Nigeria: An endogeneity corrected stochastic frontier model. Sci. Total. Environ. 745, 141151.
Orea, L., Álvarez, I.C., Jamasb, T., 2018. A spatial stochastic frontier model with omitted variables: Electricity distribution in Norway. Energy J. 39 (3), 93–116.
Peñasco, C., Romero-Jordán, D., Del Río, P., 2019. The impact of policy on the efficiency of solar energy plants in Spain: A production-frontier analysis. Econ. Energy & Environ. Policy 8 (2), 99–116.
Prokhorov, A., Tran, K.C., Tsionas, M.G., 2021. Estimation of semi- and nonparametric stochastic frontier models with endogenous regressors. Empir. Econ. 60, 3043–3068.
Sánchez-Ortiz, J., Garcia-Valderrama, T., Rodríguez-Cornejo, V., Cabrera-Monroy, F., 2021. DEA window analysis and malmquist index to assess efficiency and productivity in the spanish electricity sector. Int. J. Energy Sect. Manag. 15 (4), 765–788.
Silva, E., Macedo, P., Soares, I., 2019. Maximum entropy: A stochastic frontier approach for electricity distribution regulation. J. Regul. Econ. 55, 237–257.
Soroush, G., Cambini, C., Jamasb, T., Llorca, M., 2021. Network utilities performance and institutional quality: Evidence from the Italian electricity sector. Energy Econ. 96, 105177.
Sun, K., Wu, L., 2020. Efficiency distortion of the power generation sector under the dual regulation of price and quantity in China. Energy Econ. 86, 104675.
Tran, K.C., Tsionas, E.G., 2013. GMM estimation of stochastic frontier model with endogenous regressors. Econ. Lett. 118, 233–236.
Tran, K.C., Tsionas, E.G., 2015. Endogeneity in stochastic frontier models: Copula approach without external instruments. Econ. Lett. 133, 85–88.
Vesterberg, M., Zhou, W., Lundgren, T., 2021. Wind of change: Small-scale electricity production and distribution-grid efficiency in Sweden. Util. Policy 69, 101175.
Xu, X.-L., H.H., Chen., 2018. Examining the efficiency of biomass energy: Evidence from the Chinese recycling industry. Energy Policy 119, 77–86.
Xu, M., Tan, R., He, X., 2022. How does economic agglomeration affect energy efficiency in China?: Evidence from endogenous stochastic frontier approach. Energy Econ. 108, 105901.
Yang, C.-H., Tsou, M.-W., 2024. Evolution of technical efficiency in Chinese coal mining firms: The roles of competition and technology. Appl. Econ. 56 (2), 133–150.