Journal of Productivity Analysis (2025) 63:219–237
https://doi.org/10.1007/s11123-024-00749-9

Accounting choices in data envelopment analysis

Claire Murong Cui[1] · Julie A. Harrison[2] · Frederick Ng[2] · Paul Rouse[2]

Accepted: 21 December 2024 / Published online: 14 January 2025

Abstract

Due to the increasing availability of large-scale digitalized databases containing summarized financial accounting measures, a growing number of DEA models are using these variables. Where only accounting measures are used, we term these "FinDEA" models. Accounting measures are subject to accounting choices regarding recognition and measurement, which means that they are not necessarily equivalent to direct measures of underlying physical measures more commonly used in DEA. This paper investigates the impact of accounting choices on FinDEA results related to alternative accounting measures of capital. Using both simulated and real-world data, we find that accounting choices impact FinDEA results, with the magnitude influenced by the heterogeneity of the accounting choices and sample sizes. Our results suggest that the variations in accounting choices need to be considered as part of an assessment of the homogeneity of inputs and outputs when designing DEA models using accounting measures.

Keywords Data envelopment analysis · Firm performance · Accounting information · Accounting choices

1 Introduction

In recent years, with the growing availability of digitalized databases of financial accounting measures, these measures have increasingly been used as inputs and/or outputs in Data Envelopment Analysis (DEA) research, often in conjunction with physical measures of production inputs and outputs. In this paper, we focus on a stream of DEA research that has emerged which uses only accounting measures as inputs and outputs in DEA models of firm performance (e.g., Banker et al. 2022; Banker and Park 2021; Demerjian et al. 2012, 2013; Schwab et al. 2022).[1] We term these models "FinDEA" to distinguish them from conventional DEA models that use physical measures to calculate productivity, in line with economic productivity theories (Färe et al. 1985; Farrell 1957).

The use of accounting measures leads to a problem around potential measurement issues in FinDEA due to the existence of accounting choices. Accounting choices are choices "whose primary purpose is to influence (either in form or substance) the outputs of the accounting system" (Fields et al. 2001, p 256) and include decisions related to the choice of alternative accounting methods, estimates, or disclosures (Libby et al. 2015). While the preparation of financial statements is governed by accounting rules and standards aimed at ensuring comparability across reporting entities, different accounting choices are available under different standards. DEA, as a frontier method, is sensitive to measurement issues as it relies on extreme data points to construct a best practice frontier. Accounting choices can introduce noise and outliers that have the potential to impact the measurement of efficiency significantly.

There can be considerable variation in how FinDEA studies use accounting measures. For example, Doumpos and Cohen (2014) rejected the use of accounting variables measured using cash-based accounting systems in preference to variables measured using accrual-based accounting when designing a DEA model. They concluded that accrual accounting information incorporated a more comprehensive view of the cost-revenue structure. Rodríguez-Pérez et al. (2011) considered the impact of

✉ Claire Murong Cui
claire.mr.cui@gmail.com

✉ Paul Rouse
p.rouse@auckland.ac.nz

[1] Department of Accounting, Auckland University of Technology, Auckland, New Zealand

[2] Department of Accounting and Finance, The University of Auckland, Auckland, New Zealand

[1] We have identified 280 research studies (322 DEA models) that exclusively used accounting measures in DEA, spanning from 1990 to 2023. This list identifies DEA models developed by different research teams. The reference list is available upon request. In addition, as at September 2023, Demerjian et al. (2012) has been cited 1,698 times and Demerjian et al. (2013) has been cited 1,266 times (Google Scholar). Many of these citations relate to work where researchers have adopted the DEA measures developed by Demerjian et al. (2012, 2013).

different accounting choices on financial statement analysis, focusing on asset values prepared under historic-cost and fair-value fixed asset valuation methods. They found that these measurement methods led to substantial differences in asset values in the financial statements, significantly impacting the DEA assessment of firm performance. Further, they identified that the size of the change was not uniform across different firms or classes of assets. These papers examine accounting choices on fundamental measurement assumptions e.g., historic cost versus fair value. However, within a particular measurement choice (e.g., historic cost) and particularly for fixed and intangible assets, there are other choices concerning which accounting measure to use, which leads to the question as to whether and how differences in accounting choices will affect DEA firm performance measurement.

The problem addressed in this paper concerns accounting choices, focusing on alternative measures of capital and how accounting choices can affect these. Accounting choices related to the choice of depreciation methods, asset life estimation, and residual value estimation have been recognized as key tools that some businesses use to manage their earnings (Albonico et al. 2014; Crosby et al. 2012; Keating et al. 2000; Wilhelmsson and Roos 2024). In addition, financial statements offer researchers a choice of accounting measures related to the value and usage of capital. These include (1) gross property, plant, and plant (GPPE); (2) net property, plant, and equipment (NPPE); (3) and depreciation. GPPE and NPPE are stock measures, that is, measures of the asset value at a particular point in time. Depreciation is a flow measure, that is, a measure of asset usage over a particular period.

Accounting choices related to capital can be discretionary or non-discretionary. Prior studies have argued depreciation is a non-discretionary accounting choice that reflects the economic benefits of fixed assets over their useful economic lives. Therefore, the estimation issues related to capital are more transparent and subject to auditors' scrutiny (Burnett et al. 2012; Collins et al. 2017; Jones 1991; Lane and Willett 1997; Young 1999). Further, the literature suggests that the engagement of high-quality auditors restricts the manipulation of discretionary accruals (Alhadab and Clacher 2018). This suggests that variations in related accounting measures will be minimal.

In contrast, research has found that depreciation can be a specific discretionary accrual that management may use to manipulate financial performance (Marquardt and Wiedman 2004; Teoh et al. 1998; Vander Bauwhede et al. 2003). Researchers have argued that depreciation is commonly considered arbitrary, as it lacks an empirical reference point in the real world (Taborda and Sousa 2023; Thomas 1969, 1974). Further, the literature suggests a range of common factors, such as the complexity and difficulty in valuing infrastructure assets (Pilcher and Van Der Zahn 2010) and the inherent difficulties, arbitrariness, and ad hoc nature of depreciation (Baxter 1982), can lead to variations in accounting measures of capital across firms and asset classes.

Despite this growing literature using capital measures in FinDEA models and the variety of capital measures used, to the best of our knowledge, there has not been research that examines the impact of accounting choices on how capital is measured and its impact on FinDEA results. Accordingly, in this study, we investigate whether accounting choices related to the measurement of capital can impact the measurement of FinDEA efficiency.[2]

A simple example is where two firms use identical machinery but differ in how intensively they use their machines. Firm A uses its machines more intensively and assigns a shorter useful life resulting in a high depreciation expense and low net book value. In contrast, Firm B operates its machines less intensively and assigs a longer useful life. Consequently, depreciation is spread over a longer period with a lower annual expense and higher net book value.

These differences in accounting choices influence the efficiency rankings computed by FinDEA models. For instance, when depreciation is used as an input, Firm B appears more productive due to its lower periodic depreciation. Conversely, when gross cost is used, Firm A has higher productivity due to higher outputs as a result of more intensive production. When net book value is used, Firm A's lower net book value and higher output compared to Firm B results in a higher efficiency ranking.

This simple example indicates that FinDEA results are affected by the accounting methods chosen. These choices, including the measure of capital and assumptions about useful life, can create difference in efficiency assessments. Given the limited FinDEA research that has investigated the impact of accounting choices on FinDEA results, we examine the impact of accounting choices related to stock (e.g., net asset values) and flow (e.g., depreciation) forms of accounting measures of capital.

To test this, we use both simulated and real-world data. The simulation isolates the impact of accounting choices related to stock and flow accounting measures on FinDEA results by building incremental variations into these measures to examine their incremental impact. The empirical application uses a relatively homogeneous industry (SIC 2440 – 3412 Shipping Containers) to examine if the simulation findings hold in the real world.

[2] The focus of this paper is on FinDEA models. However, similar problems may arise in DEA models that use a mixture of accounting measures and physical measures.

Our findings from both the simulation and real-world tests indicate that accounting choices impact FinDEA results, with the magnitude influenced by heterogeneity in accounting choices and sample sizes. In particular, the empirical results align with the simulation, indicating that the simulated variation of accounting choices mirrors real-world conditions.

The findings of this study can be extended to other accounting choices such as the selection of inventory valuation methods, capital leases, revenue recognition, asset impairment and goodwill.

This study makes three main contributions. First, it provides evidence of the impact of accounting choices on the measurement of capital on FinDEA results. Researchers need to be aware that variation in accounting choices can lead FinDEA results to diverge from measures of underlying physical productivity. The study also emphasizes the inherent differences between accounting and physical measures, which can influence the homogeneity of inputs and outputs in DEA models. By addressing the variations in accounting choices, this study highlights the importance of tailoring FinDEA model designs to account for differences in accounting choices, offering a pathway for future researchers to refine their approaches in using accounting measures in DEA effectively.

Second, the simulation provides guidance on the magnitude of the variation in accounting choices and its impact on the FinDEA results, which can guide future FinDEA research regarding accounting data selection. Third, our study highlights that although accounting information is useful when measuring firm performance, accounting variables must be selected carefully as accounting choices can introduce additional attributes that can lead to a different measure than envisaged.

We proceed as follows. Section 2 discusses the literature on accounting information and the methodological considerations related to selecting variables in FinDEA. Section 3 describes the research approach using simulation. Section 4 reports the simulation results, and Section 5 describes the research approach using the empirical data. Sections 6 and 7 report the empirical results and conclusions.

2 Literature review

2.1 Methodological issues in FinDEA: selection of variables

The selection of variables in DEA plays a crucial role in determining the factors for assessing the efficiency of Decision Making Units (DMUs) (Agasisti et al. 2019; Li et al. 2021). While much of the original DEA research tends towards a productivity perspective of inputs producing outputs, there is other research that tends toward a benchmarking perspective. For example, Cook et al. (2014, p.2) note: “Although DEA has a strong link to production theory in economics, the tool is also used for benchmarking in operations management, where a set of measures is selected to benchmark the performance of manufacturing and service operations.”

From a productivity perspective, the chosen input and output factors should closely align with the factors of the underlying production process (Golany and Roll 1989; Smith 1997; Zelenyuk 2023). i.e., “the distinct kinds of goods and services used in production” (Shephard 1970, p. 13). Accounting variables may not accurately measure the production process, as all accounting measures are expressed in monetary terms (Portela and Thanassoulis 2014), aggregated at the firm level (Begen et al. 2023; Diewert and Fox 2023), and are subject to accounting choices (Rodríguez-Pérez et al. 2011).

DEA models do not necessarily have to relate to a production function and, instead, can establish a best-practice frontier (Cook et al. 2014) for benchmarking. In these models, input and output variables are selected to construct performance metrics. Mathematically, the goal is to minimize “the-less-the-better” attributes as inputs and maximize “the-more-the-better” attributes as outputs.

FinDEA studies can be found for both perspectives. For example, when selecting accounting variables according to the production process, Ren and Yang (2023) used assets, business costs, and employees to produce the value of paper products and labor services. By comparison, FinDEA models can also be used to benchmark financial efficiency using accounting variables. For instance, Demerjian et al. (2012) measured the efficiency of generating revenues by using as inputs the cost of goods sold (*COGS*), selling, general, and administrative expenses (*SG&A*), *NPPE*, operating lease, net research and development (*Net R&D*), goodwill and other intangibles. Likewise, Simper et al. (2017) measured the profit efficiency of banks by using general administration and other expenses, fees and trading expenditures, loan loss provisions and equity as inputs, non-performing loans, net interest revenue, fee and trading income, and other operating revenues as outputs.

To examine what capital-related accounting measures are used in FinDEA, we conducted a literature survey of FinDEA and identified 147 FinDEA models that included capital as an input. In this survey, NPPE was the most frequently used measure, identified in 60 models. This measure was primarily used by the literature stream based on Demerjian et al. (2012; 2013), which focuses on firm performance expressed in terms of financial efficiency (e.g., Baghdadi et al. 2018; Banker et al. 2013; Cheng et al. 2018). Demerjian et al. (2012) page 1234 explain the reason for the variable selection: “The first acquired asset, Net

Table 1 Example of accounting choices in inventory valuation

Firm A					
	Units purchased	Cost per unit	Units sold	FIFO	
				COGS	INVENTORY
First half	200	\$100	200	\$20,000 (200 × \$100)	\$30,000 (200 × \$150)
Second half	200	\$150			
Firm B					
	Units purchased	Cost per unit	Units sold	LIFO	
				COGS	INVENTORY
First half	200	\$100	200	\$30,000 (200 × \$150)	\$20,000 (200 × \$100)
Second half	200	\$150			

PP&E ("PPENT"), is reported on the balance sheet and reflects the undepreciated portion of purchased fixed assets." Subsequent research building on the pioneering work of Demerjian et al. (2012) tends to provide limited discussion regarding the rationale behind the selection of variables.

The second most used accounting variable was depreciation, which was identified in 13 models (e.g., Kapelko et al. 2014; Kapelko and Oude Lansink 2017). Kapelko et al. (2014), page 340, compared the selection of depreciation to gross investment: "Gross investments are under the control of operator who decides how much to invest, and depreciation is an administrative action."

GPPE appears to be the least frequently used measure, identified in only nine models. For instance, Yu et al. (2014) used the cost of capital (calculated using GPPE) to assess operational capability, though the selection of this variable was not explicitly discussed. The remaining 65 studies did not specify the accounting measure used.

2.2 Accounting information and accounting choices in FinDEA

Financial accounting measures consolidate business information to support stakeholder decision-making (Arya et al. 2000; Bedford 1968; Bens et al. 2018). By aggregating diverse business activities, accounting measures aggregate complex operations which can be used for DEA analysis (Banker et al. 2007). However, the inherent difference between accounting measures and physical measures, mainly relating to accounting decisions can impact the results of FinDEA (Doumpos and Cohen 2014; Rodríguez-Pérez et al. 2011).

Accounting measures that are produced according to accounting standards can be different from physical quantities because accounting measures are subject to accounting choices and estimations. A clear example can be provided by inventory valuation methods such as First-in-First-Out (FIFO) and Last-in-First-Out (LIFO). These accounting choices directly affect the reported Cost of Goods Sold (COGS) and inventory, thereby affecting the Income Statement and Balance Sheet respectively.

To illustrate, consider the example in Table 1, where Firms A and B are retailers who purchase units from wholesalers and sell them to retail customers. Both have identical purchase volumes and costs per unit for a period where the purchase cost per unit is \$100 in the first half with a large increase in cost to \$150 in the second half. Both firms have identical sales during the period. Firm A uses FIFO where it is assumed that the oldest units are sold first leaving the more recently produced units in inventory. Firm B uses LIFO where it is assumed that the most recently produced units are sold first and inventory contains the units produced earlier in the period. In periods of inflation, FIFO provides more accurate closing inventory values and less accurate cost of goods sold (COGS) while LIFO provides more accurate COGS but less accurate inventory values.[3]

Table 1 illustrates the effects of the increase in purchase costs per unit on COGS and closing inventory. Both firms purchase 400 units in the period and sell 200 units with 200 units remaining in closing inventory.

For Firm A (FIFO), the COGS is determined using the oldest inventory costs first. This leads to COGS of \$20,000, being 200 units at \$100 (the older purchase price) and the closing inventory, consisting of 200 units at \$150 each, valued at \$30,000. For Firm B (LIFO), the COGS is calculated using the cost of the most recent purchases resulting in COGS of \$30,000, being 200 units at \$150 (the most recent purchase price). The closing inventory, consisting of 200 units at \$100 each, is valued at \$20,000. Note that the physical quantities are identical for both firms.

This comparison illustrates how the accounting choice of inventory valuation influence financial measurements, reflecting variations in the cost assigned to goods sold and the valuation of unsold inventory. Companies using FIFO may present stronger profitability and efficiency in FinDEA

[3] Ling and Tinkelman (2024) observe that increased inflation in the US in 2021 and 2022 has renewed attention to inventory accounting choices, particularly the use of LIFO for tax advantages.

evaluations, if COGS is used as inputs. In contrast, companies using LIFO can generate more efficient performance if inventories are used as inputs. Ultimately, the decision between FIFO and LIFO depends on broader financial strategies (e.g. taxation) and operational insights into how inventory is managed (e.g. perishables). However, efficiency measures using FinDEA can be seriously affected depending on the accounting variables selected and the inventory valuation method used.[4]

When using accounting measures for the measurement of physical capital, measurement choices play a pivotal role in shaping accounting measures. These choices include the estimation of useful life, residual values, and the selection of depreciation methods, such as straight-line or diminishing approaches (Weygandt et al. 2015). While depreciation is often regarded as a non-discretionary accounting choice intended to systematically allocate the cost of assets over their useful economic lives (Collins et al. 2017), other researchers argue that the accounting choices for the calculation of depreciation are subject to discretion (Marquardt and Wiedman 2004). Prior studies have noted that depreciation methods are subject to managerial judgment and can sometimes be manipulated to influence financial performance (Vander Bauwhede et al. 2003). These inconsistencies highlight the tension between the objectivity of accounting standards and managerial discretion, which can impact on the accounting measures used by FinDEA models.

Some empirical studies have affirmed the usefulness of accounting information enhanced by certain conditions. McCallig et al. (2019) report that auditing opinions and credible information improves the faithful representation of accounting information. Ouda and Klischewski (2019) observed that an increase in cognitive fit between accounting information producers and users could amplify the usefulness of accounting information.

In contrast, other studies have found that contextual factors can reduce usefulness. For instance, Barth et al. (2012) argue that the comparability of accounting information is impacted by a range of factors, such as accounting standards, interpretation, auditing, and the regulatory, litigation, and enforcement environment. Cascino et al. (2021) report that while investment professionals with a firm-valuation objective consider accounting information useful, those with a managerial-performance-evaluation objective perceive accounting information as considerably less useful. Goldman and Slezak (2006) suggest that managers may manipulate accounting information in incentive contracts to deceive investors through earnings management. Hence, there is debate in the literature around the usefulness of accounting information and by implication, which accounting measures are appropriate in FinDEA models.

[4] IFRS standard IAS2 prohibits the use of LIFO whereas LIFO is allowed under US GAAP.

3 Research approach

We investigate the impact of accounting choices on FinDEA results, focusing on capital and, specifically, property, plant, and equipment (PPE), which can be proxied by alternative accounting variables. As defined by International Accounting Standard 16 (IAS 16) (IASB, 2020), PPE refers to tangible assets held for use in the production or supply of goods or services, for rental purposes, or administrative purposes, and are expected to be utilized over multiple periods.

Tangible assets form the backbone of operational and value-generation processes in asset-intensive industries like chemicals, commodity metals, semiconductors, and electric power (Kleindorfer and Wu 2003). These sectors rely extensively on physical infrastructure, such as machinery and production facilities, to sustain efficiency, ensure quality, and remain competitive in the market. For instance, manufacturers depend on cutting-edge equipment to optimize operations and achieve cost-effective production (Kleindorfer and Wu 2003). Conversely, in technology-centric or service-based industries, intangible assets, such as intellectual property, brand equity, and human capital, frequently dominate value creation (Ciftci et al. 2014).

However, the critical reliance of tangible-asset-driven sectors on physical resources highlights their importance as strategic investments. Exploring the use of tangible assets not only enables better management and optimization of these resources but also reveals insights that are foundational for sustaining efficiency and innovation in these industries.

FinDEA researchers have used various accounting variables to measure capital in production processes, including NPPE (Ahn et al. 2020; Baik et al. 2020), depreciation (Banker et al. 2020) and GPPE (Yu et al. 2014). These three accounting measures represent different aspects of PPE: (a) GPPE is a stock measure of capital, representing the purchase price for PPE at a point in time; (b) NPPE is a stock measure of capital at a current point of time, representing the net "value" that has not yet been depreciated; and (c) depreciation (DP) is a flow form of capital, representing the allocation of the depreciable value of capital to a financial year. This paper focuses on one accounting choice (estimation of useful life) and its impact on different accounting measures (GPPE, NPPE, DP) and FinDEA results.

3.1 Simple example

To demonstrate how accounting choices can influence the results of FinDEA, we use a simple example of two identical machines that differ solely in the assumed useful life. Table 2 presents two firms, A and B, both using identical machines,

Table 2 Illustrative example

Firm A				
Period	1	2	3	Productivity
No. of machines	1	2	2	
GPPE	\$100	\$200	\$200	50%
DP	\$50	\$100	\$100	100%
Accumulated DP	\$50	\$150	\$150	
NPPE	\$50	\$50	\$50	200%
SALES	\$50	\$100	\$100	
Firm B				
Period	1	2	3	Productivity
No. of machines	1	2	3	
GPPE	\$100	\$200	\$300	33%
DP	\$33	\$67	\$100	100%
Accumulated DP	\$33	\$100	\$200	
NPPE	\$67	\$100	\$100	100%
SALES	\$33	\$67	\$100	

All figures are rounded.

Productivity is Sales divided by the row variable of the steady state periods (i.e. Period 2 and 3 for Firm A, Period 3 for Firm B).

DP depreciation expense, *GPPE* gross property plant and equipment, *NPPE* net property plant and equipment, *SALES* sales revenue.

which cost \$100. Firm A's machine is subject to higher usage, and therefore, it generates \$50 in revenue per period but has a useful life (UL) of only two periods. Firm B's machine is used less and therefore generates \$33.33 in revenue per period but lasts for three periods. Focusing on the variable of interest (i.e., accounting choices regarding estimated useful life), we assume straight-line depreciation, no salvage value, and revenue evenly spread over the life of a machine. We also assume that firms can purchase only one machine per period, and each machine is removed from service (along with its gross value and accumulated depreciation) at the end of its useful life.

For example, for Firm A, one machine is owned in period 1 (GPPE = \$100) and two in period 2 (GPPE = \$200); and in period 3 the machine purchased in Period 1 is sold and another purchased, leaving two machines (GPPE = \$200). In this scenario, each type of machine reaches a steady state when the firm produces revenue of \$100 per period, which takes Firm A two periods, and Firm B three periods. In a steady state, depreciation is the same for both A and B, but the values of GPPE and NPPE differ.

If depreciation is used as the measure of capital, the usage level is captured by accounting choices such as the estimation of useful life (UL).[5] Using Sales as an output and DP as input, productivity is the same for both firms (\$100/\$100 = 100%). More generally, productivity is: $\text{Prod}_{\text{DP}} = \frac{Sales}{DP} = Sales/\frac{GPPE}{UL} = Sales\frac{UL}{GPPE}$. where straight-line depreciation $\text{DP} = \frac{GPPE}{UL}$. As useful life (UL) increases, so does Productivity and vice versa.

Using Sales as an output and GPPE as input, $\text{Prod}_{\text{GPPE}} = \frac{Sales}{GPPE}$, Firm A appears more productive (\$100/\$200 = 50%) than B (\$100/\$300 = 33%). This difference arises because Firm A's machine is used more intensively, making it more productive. This aligns with an investment perspective, indicating that for Firm A, two machines are required to generate \$100 of revenue, whereas Firm B needs three machines to achieve the same revenue.

Using Sales as an output and NPPE as input, usage and historical consumption are included being the useful life and age of the machines. In Table 2, Firm A's productivity is 200% (\$100/\$50), and Firm B is 100% (\$100/\$100). More generally, $\text{Prod}_{\text{NPPE}} = \frac{Sales}{NPPE}$ and NPPE = GPPE (1-$\left(\sum_{i=1}^{t} 1^i/UL\right)$), and $\text{Prod}_{\text{NPPE}} = Sales/\left(GPPE\left(1-\left(\frac{\sum_{i=1}^{t} 1^i}{UL}\right)\right)\right)$, where t is the current period. As UL increases (decreases), $\text{Prod}_{\text{NPPE}}$ decreases (increases). Hence, even though the machines are the same (as in this example), an accounting choice for a longer useful life will always lead to a lower productivity measurement if using NPPE.[6]

This simple example shows that when using alternative accounting measures for capital, the productivity of the two firms reflects not only differences in underlying productivity but also differences due to accounting choices, which rely on accounting assumptions and estimations. Accordingly, the choices of which accounting measure of capital and useful life can introduce variations in the FinDEA efficiency.

3.2 Base case: gross property plant and equipment (GPPE)

GPPE[7] is a stock form accounting variable as it represents the value of capital at a specific point in time and the physical capacity of capital at the date of purchase.[8]

[5] Useful life is defined as: (a) the period over which an asset is expected to be available for use by an entity; or (b) the number of production or similar units expected to be obtained from the asset by an entity (IASB, 2020).

[6] We are grateful to an anonymous reviewer who suggested the refinements to the example.

[7] GPPE is defined by IAS 16 as: "The amount of cash or cash equivalents paid, or the fair value of the other consideration given to acquire an asset at the time of its acquisition or construction or, where applicable, the amount attributed to that asset when initially recognized in accordance with the specific requirements of other IFRSs." (IASB, 2020).

[8] Under IAS 16, PPE can also be valued on a fair value basis, which "is the amount for which an asset could be exchanged between knowledgeable, willing parties in an arm's length transaction" (IASB, 2020). However, to restrict the focus of this study, only historical value is considered.

Table 3 Base case

Input distribution	Output	Inefficiency distribution
X_1: GPPE ~ N (100, 40)	$Y = (X_1 - 5)^{0.45}(X_2 - 5)^{0.45}$	N ∣(0, 0.2)∣
X_2: OPEX ~ N (100, 40)	Y inefficiency- adjusted: SALES	

GPPE (X_1) gross property, plant, and equipment, *OPEX* (X_2) operating expense, *SALES* sales, Y with inefficiency adjustment, *Y* output of Eq. (1), form Cobb-Douglas function.

Therefore, in our base case, we use GPPE as the proxy for the capital input (X_1) in a production process. We also include a second input (X_2), operating expenses (OPEX), to represent the combined factors of labor, raw materials, and other production elements. Table 3 shows GPPE and OPEX generated independently using normal distributions with a mean of 100 and a standard deviation of 40.[9] The output volume (Y) is generated from a Cobb-Douglas function (Banker and Chang 2006; Ruggiero 2007):

$$Y = (X_1 - 5)^{0.45}(X_2 - 5)^{0.45} \tag{1}$$

The exponents were selected in line with Banker and Chang (2006), who randomly generated exponents from a uniform distribution over the interval [0.4, 0.5]. To control the impact of variations in the exponents so we can focus on the impact of changes related to accounting choices, we arbitrarily selected the mid-point of 0.45 for both inputs (Harrison et al. 2012). This equates to the two inputs being equally important in the production function. Also, as the sum of the exponents is less than one this production function is in line with variable returns to scale. Both inputs are adjusted by the scalar 5 to conform to a VRS technology following Banker and Chang (2006). The inefficiency value was generated from a half-normal distribution with a mean of 0 and a standard deviation of 0.2 (Harrison et al. 2012; Ruggiero 1999). The range of inefficiencies was selected to generate a mean efficiency for each simulated sample of between 0.70 and 1.0 in line with Banker and Chang (2006). These inefficiencies were then applied to *Y* to obtain sales (SALES) adjusted for inefficiency.

We note there are limitations associated with using GPPE to proxy capital. One is that GPPE fails to capture the impact of the usage of capital over its life, and, as such, the value does not vary from year to year. Therefore, GPPE cannot fully reflect changes in the production process. However, an advantage of GPPE as a proxy for capital is that its purchase price is verifiable, and accounting choices do not usually impact its value.

[9] The simulation was designed to restrict the values to be positive.

3.3 Depreciation (DP)

Depreciation[10] arguably captures the use of capital. Depreciation is a flow form accounting variable that represents the rate at which capital is used in monetary terms. The advantage of depreciation is that it provides some information about the utilization rate of assets. However, the disadvantage of depreciation is that it can be sensitive to accounting choices, such as the depreciation method used, the estimation of an asset's useful life (UL), and the salvage value (residual amount or value). This sensitivity could potentially introduce heterogeneity into FinDEA results. In this study, we focus on one depreciation accounting method, the straight-line method, which calculates depreciation (DP) using the following formula:

$$DP = \frac{Depreciable\ Amount}{Estimated\ Useful\ Life} \tag{2}$$

Where Depreciable Amount = GPPE -- residual amount

The straight-line depreciation method estimates depreciation as the same amount each year of the PPE's estimated useful life.[11] In this study, we isolate the accounting choice of estimated useful life and do not consider the impact of accounting choices on residual values. Eq (2) shows that the depreciable amount is determined by the past transactions (GPPE and the accounting estimation of useful life, which firms determine by considering the useful life of similar assets (Weygandt et al. 2015).

Therefore, depreciation is essentially GPPE divided by the estimated useful life and in a simple technical sense, proposition one is stated as:

P_1: When depreciation is used to calculate the FinDEA efficiency score, a longer estimated useful life generates a smaller amount of depreciation for a year and, thus, a higher FinDEA efficiency score for a particular DMU.[12]

[10] Depreciation is defined by IAS 16 as: "The systematic allocation of the depreciable amount of an asset over its useful life" (IASB, 2020).

[11] Another common depreciation method is the diminishing value method, which depreciates a fixed percentage of the previous year's closing balance. For this amount, although the deprecation rate is fixed, it applies to a diminishing closing balance each year. This study uses the straight-line depreciation method since it is the method used by the majority of the firms in the shipping container industry, which we used in our empirical analysis.

[12] The proposition follows the fundamental axiom of Monotonicity in DEA (Banker 1993; Banker et al. 1984; Shephard 1970).

Table 4 Proposition one: depreciation expense

Scenario	Input1	Input2	Output
Constant	DP = GPPE/UL	OPEX ~ N(100, 40)	SALES
	UL = 10; AGE = 5.3		
Narrow	DP = GPPE/UL	OPEX ~ N(100, 40)	SALES
	UL = 5.3/ULC		
	ULC ~ N (0.53, 0.0265)		
Broad	DP = GPPE/UL	OPEX ~ N(100, 40)	SALES
	UL = 5.3/ULC		
	ULC ~ N (0.53, 0.1325)		

AGE age of capital, *DP* depreciation expense, *GPPE* gross property plant and equipment, *OPEX* operating expense, *SALES* sales; *UL* useful life of capital, *ULC* useful life consumed, the proportion of the age of capital to the useful life of capital.

3.4 Net property plant and equipment (NPPE)

NPPE[13] is a stock form accounting variable, calculated as the value of the capital at the time of purchase (GPPE), less the cost of the accumulated depreciation. It represents a net capital value at the end of a financial year.

The advantages of NPPE are that it considers both the cost of the asset and the accumulated usage over past years. However, NPPE reflects the remaining non-depreciated portion of the asset rather than the value used in the current production period. Additionally, NPPE is highly sensitive to accounting choices, as its calculation is influenced by choices made regarding depreciation (DP) and the age of capital (AGE), as shown in Eq. (3):

$$NPPE = GPPE - AGE \times DP \tag{3}$$

To isolate the accounting choices that impact the value of NPPE, Eqs. (2) and (3) are combined in Eq. (4) and $\frac{AGE}{UL}$ is named useful life consumed (ULC).

$$NPPE = GPPE \times (1 - \mathrm{ULC}) \tag{4}$$

GPPE is determined by a past transaction (price at the time of purchase), and NPPE is a function of the ULC. Hence, proposition two, again in a simple technical sense, is stated as follows:

P_2: When NPPE is used to calculate FinDEA efficiency scores, a higher ULC ($\frac{AGE}{UL}$) generates a lower NPPE for a year and a higher FinDEA efficiency score for a particular DMU. [14]

[13] NPPE is defined by the IAS 16 as: "The amount at which an asset is recognized after deducting any accumulated depreciation and accumulated impairment losses". (IASB, 2020)

[14] The proposition follows the fundamental axiom of Monotonicity in DEA (Banker 1993; Banker et al. 1984; Shephard 1970).

3.4.1 Scenarios

Three scenarios were developed to examine the impact of the variability of accounting choices on FinDEA results: constant, narrow, and broad, where incremental levels of heterogeneity were introduced using different accounting choices.

Useful life consumed (ULC) is the key parameter that varies across the three scenarios and is used to generate NPPE from Eq. (4). Depreciation (DP) is solely a function of useful life. However, it is necessary to use ULC to ensure that estimates of capital age (AGE) are always less than the estimated useful life for a particular PPE.

The simulated ULC was distributed around a mean of 53% to ensure the simulated and empirical tests were comparable. The GPPE, OPEX, and SALES values generated by Eq. (1) were consistently used in all the simulated FinDEA models using one iteration for all three scenarios. For proposition P_1, which examines the impact of estimated useful life (UL) on depreciation, in the Constant Scenario, AGE is 5.3, and UL is 10 (Table 4). In the narrow and Broad Scenarios, UL incorporates variations generated using a randomized ULC. We used the coefficient of variation (CV) to determine the standard deviations for the narrow and broad case. The CV for ULC was 0.05 for the Narrow Scenarios and 0.25 for the Broad Scenario; the selection of 0.05 for the narrow and 0.25 for the Broad Scenarios was based on a review of empirical FinDEA literature (Demerjian 2018; Demerjian et al. 2012, 2013). This review identified that the mean CV for different industries varied from 0.06 to 0.91, with an average of 0.35. In the Narrow (Broad) Scenario, the ULC was generated from a normal distribution with a mean of 0.53 and a standard deviation of 0.0265 (0.1325). Using this ULC, UL was calculated as 5.3/ULC, assuming that the AGE is 5.3 and DP was calculated from Eq. (2).

Proposition P_2 follows the same simulation approach as described in the previous paragraph. Table 5 sets out the equations showing the impact of ULC on NPPE. In the Constant Scenario, ULC is 0.53, AGE is 5.3, and UL is 10. In the narrow and Broad Scenario, ULC varies according to the standard deviations for each scenario being 0.0265 and 0.1325 respectively, with NPPE calculated using Eq. (4) as shown in the table.

3.4.2 DEA models

Three sample sizes were chosen ($N = 24$, 96, 384) and repeated 1000 times using the Monte Carlo method (Andor and Hesse 2014; Banker et al. 1993; Ruggiero 2007). The minimum sample size was decided following various 'rules of thumb' (Golany and Roll 1989; Khezrimotlagh et al.

Table 5 Proposition two: net property plant and equipment

Scenario	Input1	Input2	Output
Constant	NPPE = GPPE × (1 − ULC)	OPEX ~ N (100, 40)	SALES
	AGE = 5.3		
	UL = 10		
Narrow	NPPE = GPPE × (1 − ULC)	OPEX ~ N (100, 40)	SALES
	UL = 5.3 ÷ *PD*		
	ULC ($\frac{AGE}{UL}$) ~ N (0.53, 0.0265)		
Broad	NPPE = GPPE × (1 − ULC)	OPEX ~ N (100, 40)	SALES
	UL = 5.3 ÷ *PD*		
	ULC ($\frac{AGE}{UL}$) ~ N (0.53, 0.1325)		

AGE age of capital, *GPPE* gross property plant and equipment, *NPPE* net property plant and equipment, *OPEX* operating expense, *SALES* sales; *UL* useful life of capital, *ULC* useful life consumed, the proportion of the age of capital to the useful life of capital.

2021; Podinovski and Thanassoulis 2007) and prior literature (Ruggiero 1999).

The FinDEA models applied a Variable Returns to Scale (VRS) production function (Charnes et al. 1978), where the sum of the exponents in the Cobb-Douglas function is 0.9, which is less than one, suggesting a Decreasing Returns to Scale (DRS) (Oh and Shin 2015). Furthermore, an output orientation was chosen to match the Cobb-Douglas function. Finally, to avoid any effects due to the randomization process, each scenario was repeated 1000 times (Jradi and Ruggiero 2019; Ruggiero 2007).

4 Simulation results

4.1 Descriptive statistics

Table 6 reports the descriptive statistics for the variables used in the simulation with iterations of 1000. The distribution of variables aligns with the specified parameters in the simulation design in Tables 4 and 5. For instance, the inputs (GPPE, OPEX) in the base case have an approximate mean of 100 and a standard deviation of approximately 40.

4.2 Efficiency scores

Table 7 presents the efficiency scores from the FinDEA model for sample sizes of 24 and 384.[15] Fig. 1 shows boxplots illustrating the distribution of efficiency scores for the sample size 384 and iteration of 1000. First, the findings from the efficiency scores indicate that when the accounting choices remain constant, the FinDEA results are the same regardless of whether capital is measured using GPPE, DP, or NPPE. When accounting choices vary, the FinDEA results using DP as a measure of capital stay closer to the base case (GPPE) compared to results using NPPE. As the variation in accounting choices increases, from the narrow to the Broad Scenarios, the FinDEA results generated by both DP and NPPE diverge further from the base case (GPPE).

In Table 7, Panel B ($N = 384$), the narrow mean efficiency score generated by FinDEA using DP is 0.8630, whereas the mean efficiency score using NPPE is 0.8599. Additionally, DP identifies more efficient DMUs (7.75%) than NPPE (7.52%). Moreover, the range of efficiency scores generated with DP (0.9427) is narrower than that of NPPE (0.9487). The trend can be seen in Fig. 1 ($N = 384$) where the boxplot generated by DP generally sits closer to the full efficiency (efficiency score = 1) than the boxplot generated by NPPE.

Second, the variation of FinDEA results increases moving from the constant to the Broad Scenario. Consequently, the efficiency scores for the DMUs calculated under the Broad Scenarios exhibit higher variability (SD and CV) than the Narrow Scenario. For instance, in Table 7, Panel B, in the Broad Scenario, the FinDEA results calculated using DP have a SD (CV) of 0.1320 (0.1753), whereas the Narrow Scenario using DP has a SD of 0.1016 (0.1177).

Third, when the sample size increases from Panel A ($N = 24$) to Panel B ($N = 384$), there is a decline in the average efficiency score, which aligns with prior studies (Charles et al. 2019; Diewert and Mendoza 1995; Dyson et al. 2001). Also, the percentage of efficient DMUs in the Narrow Scenario changes from 41.87% in Panel A to 7.75% in Panel B for DP. These findings are shown in Fig. 1 ($N = 384$).

4.3 FinDEA results compared to the base case

Table 8 presents the correlations and the mean absolute difference (MAD) between the FinDEA results for each scenario compared to the base case (GPPE). Overall, the findings confirm Table 7 that accounting variables yield the same FinDEA efficiency score when the accounting choices are constant. However, in practice, accounting choices are likely to differ among firms, and the test shows increasing variability from the constant to narrow and Broad Scenarios.

First, FinDEA results generated using either DP or NPPE diverge from the base case when there is variation in accounting choices, but NPPE is more sensitive than DP. Table 8 demonstrates a consistent pattern where the FinDEA efficiency scores generated using DP (P_1) are more

[15] In order to reduce the paper length only the results for these sample sizes are reported. Similar results were found for N = 96.

Table 6 Descriptive statistics for simulated data

Panel A: $N = 24$, iteration = 1000

	Scenarios	Variables	MEAN	SD	CV	MIN	Q1	Median	Q3	MAX
Base		*GPPE (X1)*	100.2343	39.9312	0.3984	0.1737	73.1368	100.1057	127.0364	250.1007
		OPEX (X2)	100.1110	39.7447	0.3970	0.1682	72.9124	100.2994	127.0780	265.5155
		Y	57.8301	17.1204	0.2960	0.8032	46.6398	58.4257	69.5237	125.9291
		Inefficiency	0.1596	0.1196	0.7497	0.0000	0.0641	0.1351	0.2296	0.8169
		SALES (Adj-Y)	49.6270	15.8221	0.3188	0.7814	38.8331	49.5425	60.2651	113.7349
Factors	Constant	*UL*	10.00	0.00	0.00	10.00	10.00	10.00	10.00	10.00
		ULC	0.53	0.00	0.00	0.53	0.53	0.53	0.53	0.53
	Narrow	*UL*	10.0246	0.5070	0.0506	8.3644	9.6740	9.9988	10.3523	12.4460
		ULC	0.5300	0.0266	0.0502	0.4258	0.5120	0.5301	0.5479	0.6336
	Broad	*UL*	10.7875	3.6048	0.3342	5.3887	8.5459	9.9858	12.0070	50.0000
		ULC	0.5303	0.1320	0.2489	0.0490	0.4414	0.5308	0.6202	0.9835
P1	Constant	*DP*	10.0238	3.9968	0.3987	0.0174	7.3136	10.0109	12.7036	25.0101
	Narrow	*DP*	10.0242	4.0296	0.4020	0.0169	7.2845	9.9707	12.7187	26.0211
	Broad	*DP*	10.0228	4.7917	0.4781	0.0169	6.5903	9.6145	12.9810	38.5243
P2	Constant	*NPPE*	47.1118	18.7754	0.3985	0.0816	34.3744	47.0527	59.7071	117.5473
	Narrow	*NPPE*	47.1060	18.9879	0.4031	0.0843	34.1041	46.9925	59.7371	124.9452
	Broad	*NPPE*	47.1150	23.5424	0.4997	0.0715	30.1918	44.6718	61.4024	179.4165

Panel B: $N = 384$, iteration = 1000

	Scenarios	Variables	MEAN	SD	CV	MIN	Q1	Median	Q3	MAX
Base		*GPPE (X1)*	100.0828	39.7994	0.3977	0.0015	72.9257	100.0270	127.0116	283.7432
		OPEX (X2)	100.0336	39.6664	0.3965	0.0060	73.0001	99.9963	126.9716	278.1787
		Y	57.8631	16.9356	0.2927	0.2812	46.7039	58.3169	69.4621	127.7804
		Inefficiency	0.1598	0.1207	0.7556	0.0000	0.0638	0.1349	0.2307	0.9742
		SALES (Adj-Y)	49.6693	15.7017	0.3161	0.2176	38.9848	49.4802	60.2350	121.1245
Factors	Constant	*UL*	10.00	0.00	0.00	10.00	10.00	10.00	10.00	10.00
		ULC	0.53	0.00	0.00	0.53	0.53	0.53	0.53	0.53
	Narrow	*UL*	10.0254	0.5048	0.0503	8.1462	9.6734	10.0017	10.3497	13.1558
		ULC	0.5300	0.0265	0.0500	0.4029	0.5121	0.5299	0.5479	0.6506
	Broad	*UL*	10.8033	3.6187	0.3350	5.3026	8.5608	10.0080	12.0340	50.0000
		ULC	0.5298	0.1325	0.2502	0.0025	0.4404	0.5296	0.6191	0.9995
P1	Constant	*DP*	10.0083	3.9799	0.3977	0.0001	7.2926	10.0027	12.7012	28.3743
	Narrow	*DP*	10.0082	4.0171	0.4014	0.0002	7.2597	9.9772	12.7043	28.3838
	Broad	*DP*	10.0058	4.8072	0.4804	0.0001	6.5671	9.5434	12.9664	39.4617
P2	Constant	*NPPE*	47.0389	18.7057	0.3977	0.0007	34.2751	47.0127	59.6954	133.3593
	Narrow	*NPPE*	47.0394	18.9169	0.4022	0.0006	34.1071	46.8826	59.7311	139.3684
	Broad	*NPPE*	47.0538	23.5219	0.4999	0.0008	30.1229	44.5742	61.3123	198.6557

The descriptive features include the average value (MEAN), standard deviation (SD), the coefficient of variation, which is the ratio of the standard deviation to the mean (CV), minimal value (MIN), 25th percentile (Q1), 50th percentile (Median), 75th percentile (Q3), and the maximum value (MAX).

All efficiency scores are generated from the VRS output orientation model.

The descriptive statistics for $N = 96$, i = 1000, are reported in Supplementary Appendix A.

DP depreciation expense; *GPPE* (X1) gross property plant and equipment; Inefficiency: inefficiency value was generated from a half-normal distribution with a mean of 0 and a standard deviation of 0.2 (Table 3) *NPPE* net property plant and equipment, *OPEX* (X2) operating expense, *SALES* sales, *UL* useful life of capital, *ULC* useful life consumed, the proportion of the age of capital to the useful life of capital, *Y* output of Eq. (1), form Cobb-Douglas function.

Table 7 Descriptive statistics for simulated efficiency scores

Panel A: $N = 24$, iteration = 1000										
	Variables	MEAN	SD	CV	100%	Range	MIN	Q1	Median	Q3
Base case	GPPE	0.9317	0.0901	0.0967	0.4284	0.5252	0.4748	0.8817	0.9786	1.0000
Constant	DP (P1)	0.9317	0.0901	0.0967	0.4283	0.5252	0.4748	0.8817	0.9786	1.0000
	NPPE (P2)	0.9317	0.0901	0.0967	0.4281	0.5252	0.4748	0.8817	0.9786	1.0000
Narrow	DP (P1)	0.9291	0.0920	0.0990	0.4187	0.5558	0.4442	0.8773	0.9749	1.0000
	NPPE (P2)	0.9284	0.0924	0.0995	0.4195	0.5360	0.4640	0.8754	0.9735	1.0000
Broad	DP (P1)	0.8918	0.1205	0.1351	0.3553	0.6654	0.3346	0.8094	0.9302	1.0000
	NPPE (P2)	0.8852	0.1253	0.1416	0.3475	0.6620	0.3380	0.7990	0.9216	1.0000
Panel B: $N = 384$, iteration = 1000										
	Variables	MEAN	SD	CV	100%	Range	MIN	Q1	Median	Q3
Base case	GPPE	0.8773	0.0999	0.1139	0.0977	0.9466	0.0534	0.8120	0.8945	0.9614
Constant	DP (P1)	0.8773	0.0999	0.1139	0.0976	0.9466	0.0534	0.8120	0.8945	0.9614
	NPPE (P2)	0.8773	0.0999	0.1139	0.0976	0.9466	0.0534	0.8120	0.8945	0.9614
Narrow	DP (P1)	0.8630	0.1016	0.1177	0.0775	0.9427	0.0573	0.7960	0.8777	0.9446
	NPPE (P2)	0.8599	0.1021	0.1187	0.0752	0.9487	0.0513	0.7926	0.8743	0.9411
Broad	DP (P1)	0.7528	0.1320	0.1753	0.0452	0.9578	0.0422	0.6605	0.7524	0.8458
	NPPE (P2)	0.7381	0.1392	0.1886	0.0426	0.9632	0.0368	0.6421	0.7382	0.8359

(1) The descriptive features include the mean value (MEAN), standard deviation (SD), the coefficient of variation (CV), that is, the ratio of the standard deviation to the mean, the proportion of efficient DMUs (100%), the proportion of DMUs that are at least 99% efficient (99%), the range of efficiency scores (Range), the minimal value (MIN), 25th percentile (Q1), 50th percentile (Median), and 75th percentile (Q3).

All efficiency scores are generated from the VRS output orientation model.

The results for $N = 96$ are reported in Supplementary Appendix B.

DP depreciation expense, *NPPE* net property plant and equipment.

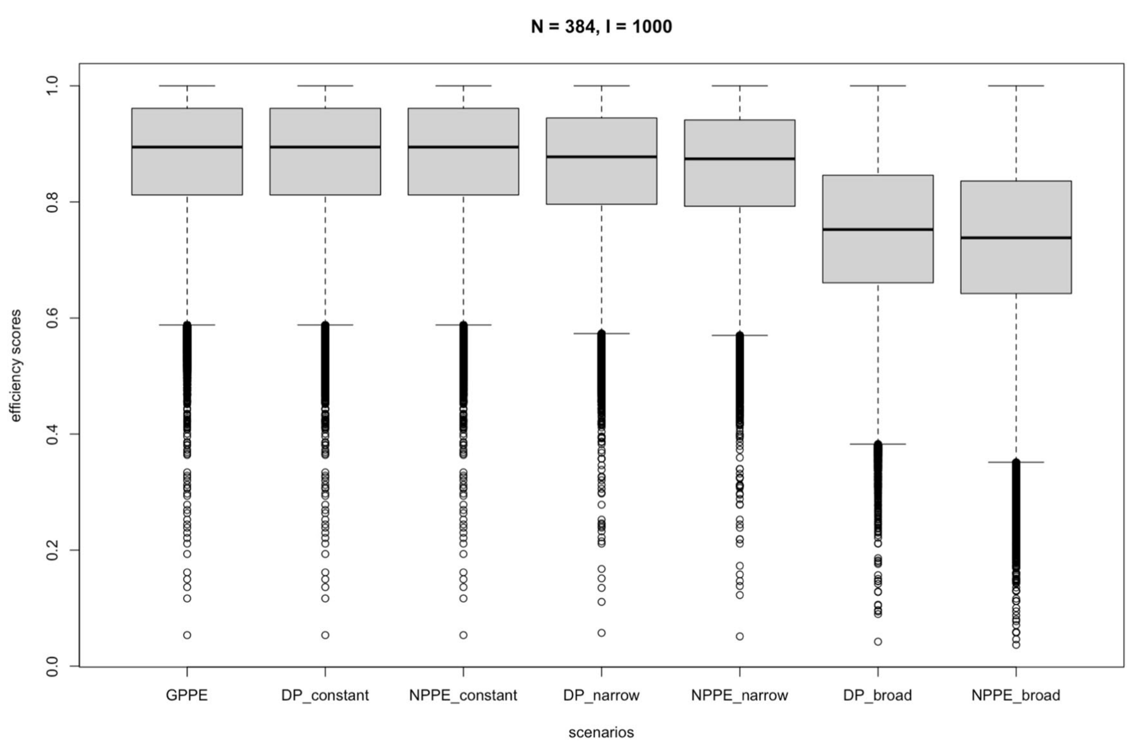

Fig. 1 Box plot for the distribution of simulated efficiency scores for $N = 384$ with 1000 iterations

Table 8 Variation of financial DEA results with simulated data

Panel A: $N = 24$, iteration = 1000									
P1	Constant			Narrow			Broad		
	Pearson	Spearman's	MAD	Pearson	Spearman's	MAD	Pearson	Spearman's	MAD
	1	1	0.0000	0.9718	0.8817	0.0119	0.6966	0.6314	0.0639
	(0.0000***)	(0.0000***)		(0.0000***)	(0.0000***)		(0.0122**)	(0.0195**)	
P2	1	1	0.0000	0.9665	0.8721	0.0130	0.6517	0.5900	0.0690
	(0.0000***)	(0.0000***)		(0.0000***)	(0.0000***)		(0.0216**)	(0.0321**)	
Panel B: $N = 384$, iteration = 1000									
P1	Constant			Narrow			Broad		
	Pearson	Spearman's	MAD	Pearson	Spearman's	MAD	Pearson	Spearman's	MAD
	1	1	0.0000	0.9748	0.9655	0.0203	0.6543	0.6337	0.1309
	(0.0000***)	(0.0000***)		(0.0000***)	(0.0000***)		(0.0000***)	(0.0000***)	
P2	1	1	0.0000	0.9683	0.9576	0.0235	0.6001	0.5855	0.1412
	(0.0000***)	(0.0000***)		(0.0000***)	(0.0000***)		(0.0000***)	(0.0000***)	

DEA models are the Variable Returns to Scale, output orientation.

The criteria for FinDEA results are the Pearson Correlation (Pearson), the Spearman's Ranking Correlation (Spearman's), and the Mean Absolute Deviation (MAD).

A correlation coefficient greater than 0.8 is considered a moderate level of correlation; a coefficient greater than 0.9 indicates a high level of correlation.

The comparison point is the base case where FinDEA results are generated using *GPPE* gross property plant, and equipment.

The results of $N = 96$ are reported in the Supplementary Appendix C.

DP depreciation expense, *NPPE* net property plant and equipment.

The numbers in parentheses are *P*-values. *** represents for significant level of < 0.01, ** represents for significant level of significant level of < 0.05, * represents for significant level of < 0.1.

closely correlated with the scores generated by GPPE (the base case) than the FinDEA efficiency scores generated by NPPE (P_2).

Second, it is observed that greater variability of the useful life consumed (ULC), i.e. moving from the narrow to Broad Scenarios, leads to more divergence from the base case. This observation is supported by lower correlation coefficients, less significant *p*-values, and larger MAD values. For example, in Table 8, Panel A ($N = 24$), for DP (P_1), the correlation and MAD values in the Constant Scenario exhibit near-perfect correlation with the base case (correlation coefficient of 1, *p*-values of 0.0000, and MAD of 0.0000). However, in the Narrow Scenario, the Pearson correlation coefficients decrease to 0.9718, with a significant *p*-value (0.0000). In the Broad Scenario, the Pearson correlation coefficient decreases to approximately 0.6966, with a significant *p*-value (0.0122) These patterns can also be observed in the results obtained using NPPE (P_2) and across different sample sizes.

Third, as the sample sizes increase from $N = 24$ to $N = 384$, in the Narrow Scenario, the correlations for DP (P_1) and NPPE (P_2) compared to GPPE improve. For example, in the Narrow Scenario (Table 8), compared with Panel B ($N = 384$), Panel A ($N = 24$) shows relatively lower correlation coefficients between NPPE (P_2) and GPPE, with Pearson (Spearman's) correlation coefficients of 0.9665 (0.8721) and *p*-values of 0.0000 (0.0000).

However, in Panel B ($N = 384$), in the Narrow Scenario, the correlation coefficients between NPPE (P_2) and GPPE improve to 0.9683 (0.9576) with same highly significant *p*-values of 0.0000.

Fourth, the MAD increases with sample size, indicating the FinDEA efficiency scores with DP (P_1) and NPPE (P_2) diverge further from the base case (GPPE). In Table 8, Narrow Scenario, the MAD when the sample size is 24 (Panel A) is 0.0119 between the results with DP and the base case (GPPE). However, when the sample size is 384 (Panel B), the MAD increases to 0.0203. The trend of MAD contradicts the correlations since increasing the sample size does not converge the results to the base case. This finding is consistent with previous studies showing an increase in MAD with sample size when there are larger measurement errors (Banker et al. 1993; Ruggiero 2004).

5 Empirical research approach

One of the limitations of simulation studies is the potential for unrealistic assumptions. For instance, the simulated test only focuses on one accounting choice, the estimated UL.

However, in practice, other accounting choices (e.g., depreciation method) and operational characteristics (e.g., industry and machine features) can also impact accounting measures used in FinDEA. The simulated test employed artificial data that relied on specific assumptions, parametric functions, and distributional properties. As a result, the generalizability of the results may be limited, as the artificial data may not accurately represent real-world conditions (Harrison et al. 2012). To address this limitation, an empirical test was conducted to investigate whether similar results were observed using real-world data.

Empirical data were obtained from the COMPUSTAT database for a five-year period. A single industry was selected to ensure an adequate sample size while minimizing the potential impact of significant technological changes (Färe et al. 1994; Golany and Roll 1989). The sample chosen was from the shipping container industry (SIC 2440 - 3412 Shipping Containers) based on the Fama-French 48 industry classification (Demerjian et al. 2012; Fama and French 1997). This industry was assessed as relatively homogeneous based on the mean efficiency score of 0.934 reported by Demerjian (2018).

Table 9 Sample description

Industry	Shipping container
All COMPUSTAT observations from 2015–2019	59
Delete observations with missing values	2
Delete observations with zero values	−0
Delete extreme values	−0
Full sample	57

All values are inflated by the consumer price index (CPI) for the financial year 2019.

We checked the observations with missing and zero values, following the approach by Golany and Roll (1989), and observations with extreme values (estimated UL and AGE over 100 years) were also examined (there are zero cases in the sample set). The final sample (Table 9) comprised 57 DMUs. To account for inflation, all values were adjusted using the consumer price index (CPI) to reflect the values for the 2019 financial year (Rouse and Tripe 2016). To improve the robustness of the result, without expanding the time frame and exposing the sample to significant technological change, a sub-sample of 30 DMUs was randomly resampled with replacements 1000 times (Efron 1993).

The accounting variables used in the FinDEA models for the empirical test were the same as those used in the simulated test. OPEX measured labor and other non-capital expenses. Depending on the model, capital was measured using GPPE, NPPE, or DP. The output was sales (SALES). Both constant returns to scale (CRS) and variable returns to scale (VRS) models were analyzed (Banker et al. 1984).

6 Empirical results

6.1 Descriptive statistics

Table 10 presents the descriptive statistics for the empirical data. The variables used are the same as in the simulated test, except for UL, AGE, and ULC, which are not publicly available and were estimated. The assumptions made for the estimation are consistent with the simulation: all assets (PPE) have zero residual values, and the straight-line depreciation method is used. Hence, UL is calculated from Eq. (2), and AGE is calculated from Eq. (3).

Table 10 Descriptive statistics for raw empirical data

Shipping Container Industry FY 2015–2019 ($N = 57$)								
	MEAN	SD	CV	MIN	Q1	MED	Q3	MAX
GPPE	5935.88	3550.00	0.60	700.11	3757.26	5717.92	7253.70	17461.30
DP	420.02	307.48	0.73	39.05	206.48	386.53	490.00	1440.50
NPPE	3132.78	2522.21	0.81	275.56	1339.47	2968.90	3456.60	11189.50
OPEX	5991.54	3261.59	0.54	711.38	3883.60	5398.58	7762.38	15148.10
SALES	7116.49	3913.50	0.55	785.60	4595.00	6660.00	9219.88	18289.00
UL	15.85	4.26	0.27	6.72	13.37	16.22	17.94	28.01
AGE	8.34	3.19	0.38	2.61	5.88	8.70	10.43	14.89
ULC	0.51	0.11	0.21	0.24	0.44	0.53	0.58	0.68

All variables, except the UL, AGE, and ULC, are retrieved from the COMPUSTAT database in millions of United States dollars (USD).

The values of UL, AGE, and ULC are estimates based on accounting definitions, assuming zero residual value and straight-line depreciation.

AGE age of capital, *DP* depreciation expense, *GPPE* gross property plant and equipment, *NPPE* net property plant and equipment, *OPEX* operating expense, *SALES* sales, *UL* useful life of capital, *ULC* useful life consumed, the proportion of the age of capital to the useful life of capital.

Table 11 Descriptive statistics for bootstrapped empirical data

Shipping Container Industry ($N = 30$, iteration = 1000)								
	MEAN	SD	CV	MIN	Q1	MED	Q3	MAX
GPPE	5939.74	3547.40	0.60	700.11	3757.26	5717.92	7253.70	17461.30
DP	420.21	307.09	0.73	39.05	206.48	386.53	490.00	1440.50
NPPE	3135.67	2517.22	0.80	275.56	1339.47	2968.90	3456.60	11189.50
OPEX	5981.42	3253.53	0.54	711.38	3883.60	5398.58	7762.38	15148.10
SALES	7106.15	3904.33	0.55	785.60	4595.00	6660.00	9219.88	18289.00
UL	15.85	4.20	0.27	6.72	13.37	16.22	17.94	28.01
AGE	8.35	3.16	0.38	2.61	5.88	8.70	10.43	14.89
ULC	0.51	0.11	0.21	0.24	0.44	0.53	0.58	0.68

All variables, except for the UL, AGE, and ULC, are retrieved from the COMPUSTAT database and in the units of millions of United States dollars (USD).

The values of UL, AGE, and ULC are estimates based on accounting definitions, assuming zero residual value and straight-line calculation method.

AGE age of capital, *DP* depreciation expense, *GPPE* gross property plant and equipment, *NPPE* net property plant and equipment, *OPEX* operating expense, *SALES* sales, *UL* useful life of capital, *ULC* useful life consumed, the proportion of the age of capital to the useful life of capital.

Table 12 Descriptive statistics bootstrapped efficiency scores

Panel A: VRS Input-oriented ($N = 30$, iteration = 1000)								
	MEAN	SD	CV	100%	Range	MIN	Q1	MED
GPPE	0.9699	0.0244	0.0251	7.05	0.0828	0.9172	0.9484	0.9715
DP	0.9696	0.0244	0.0251	7.28	0.092	0.908	0.9502	0.9699
NPPE	0.9782	0.0217	0.0222	9.56	0.0759	0.9241	0.9605	0.9818
Panel B: VRS Output-oriented ($N = 30$, iteration = 1000)								
	MEAN	SD	CV	100%	Range	MIN	Q1	MED
GPPE	0.9694	0.0245	0.0252	7.05	0.0793	0.9207	0.9475	0.9703
DP	0.97	0.0237	0.0244	7.28	0.0865	0.9135	0.95	0.9701
NPPE	0.9782	0.0216	0.022	9.56	0.0749	0.9251	0.9606	0.9812
Panel C: CRS ($N = 30$, iteration = 1000)								
	MEAN	SD	CV	100%	Range	MIN	Q1	MED
GPPE	0.9536	0.027	0.0283	2.56	0.1194	0.8806	0.9375	0.9502
DP	0.9491	0.0268	0.0282	2.58	0.1205	0.8795	0.9356	0.9462
NPPE	0.9623	0.0231	0.024	2.49	0.0907	0.9093	0.9462	0.9623

For each iteration, the sample size is 30, repeated 1000 times.

The descriptive features include the average value (MEAN), standard deviation (SD), the coefficient of variation, that is the ratio of the standard deviation to the mean (CV), the portion of efficient DMUs (100%), the range of efficiency scores (Range), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3).

DP depreciation expense, *GPPE* gross property plant and equipment, *NPPE* net property plant and equipment.

Table 10 indicates that the empirical data represent a reasonably homogeneous sample set with the coefficient of variation (CV) for all variables ranging from 0.2 to 0.8, which are relatively small compared to the other industries, as evidenced by the literature (Demerjian 2018; Demerjian et al. 2012, 2013). The accounting choices, such as UL and ULC, have CVs ranging from 0.2 to 0.38 and are relatively homogeneous.

Table 11 presents the descriptive statistics of the bootstrapped data. The resampling process retains the key characteristics of the raw sample presented in Table 10. For instance, the mean difference of OPEX between the raw data (5991.54) and the bootstrapped data (5981.42) is 10.12, less than one percent. Furthermore, the coefficient of variation values remain unchanged between the raw data and the bootstrapped data for all variables.

6.2 Efficiency scores

Table 12 reports the efficiency scores. The VRS input-oriented results (Panel A) and VRS output-oriented

Table 13 Variation of FinDEA results with bootstrapped empirical data

Shipping Container Industry ($N = 30$, iteration = 1000)									
	VRS input-orientation			VRS output-orientation			CRS		
	Pearson	Spearman's	MAD	Pearson	Spearman's	MAD	Pearson	Spearman's	MAD
DP	0.7315 (0.0007***)	0.7245 (0.0008***)	0.0118	0.7333 (0.0009***)	0.7290 (0.0009***)	0.0116	0.6506 (0.0048***)	0.6057 (0.0112**)	0.0160
NPPE	0.6610 (0.0017***)	0.6579 (0.0014***)	0.0137	0.6527 (0.0019***)	0.6512 (0.0016***)	0.0138	0.6046 (0.0028***)	0.5721 (0.0050***)	0.0184

The criteria for the variation of FinDEA results are the Pearson Correlation (Pearson), the Spearman's Ranking Correlation (Spearman's), and the Mean Absolute Deviation (MAD).

The base case is the FinDEA results generated using GPPE.

DP depreciation expense, *NPPE* net property plant and equipment.

The numbers in parentheses are *p*-values. *** represents for significant level of < 0.01, ** represents for significant level of significant level of < 0.05, * represents for significant level of < 0.1.

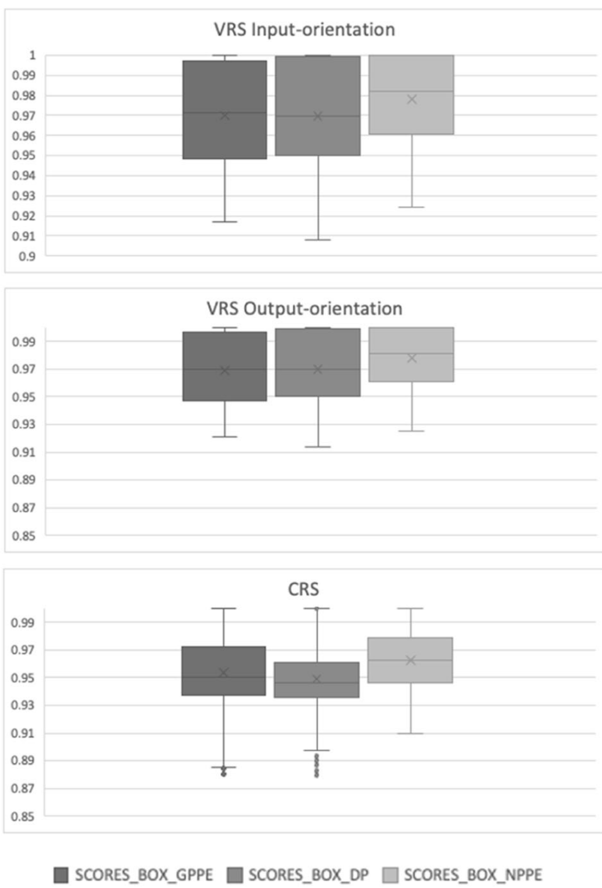

Fig. 2 Box plot for the distribution of bootstrapped efficiency scores

results (Panel B) are very similar. The mean efficiencies are consistently high, sitting above 0.90, and the standard deviation (coefficient of variation) is approximately 0.02 (0.02), aligning with the expected level of homogeneity found in the prior literature (Demerjian 2018).

6.3 FinDEA results compared to the simulation results

Table 13 presents the correlations and MAD between the DP and NPPE results and the base case (GPPE). Overall, the shipping container industry exhibits similar results to the simulation (Table 8). The correlations align similarly, falling between the narrow and broad ranges outlined in Table 8, given the comparable coefficients of variation (CVs) between the simulation and the empirical application. This confirms that the simulation design is reasonable for exploring and explaining real-world phenomena.

Figure 2 shows that the FinDEA results generated using DP were closer to the results for GPPE compared to the results using NPPE. This pattern also follows the results found when using simulated data (Fig. 1). For example, in Table 13, VRS input-orientation, the Pearson (Spearman's) correlation coefficient for DP is 0.7315 (0.7245), which is higher than the coefficient for NPPE, which is 0.6610 (0.6579). Also, the MAD of the efficiency scores generated with DP is 0.0118, which is slightly closer to the base case than the MAD generated by NPPE (0.0137).

In practice, the observed differences within the shipping container industry between accounting variables can be attributed to specific characteristics of the shipping container industry. The shipping container industry provides shipping services and packaging goods for customers worldwide. The primary equipment used in this industry is container boxes made of steel, which have relatively short useful lives and are regularly updated to ensure security during transportation (Wan et al. 2016). The shipping container industry has experienced significant growth in the past 50 years, primarily driven by the expansion of international trade (Lee et al. 2014). Given the nature of the international market and the relatively recent establishment of many companies in this industry, the GPPE value in the shipping container industry is likely to be relatively homogeneous.

Additionally, DP may be less homogeneous since the usefulness of containers varies depending on the types of goods being shipped and the length of the travel route. Containers exposed to harsh traveling conditions are frequently replaced (Lee et al. 2014). Further, NPPE may also be less homogeneous in the shipping container industry because it varies regarding useful life and asset age. Factors such as the length of time in business, changes in customer relationships, and the durability of the box material can contribute to the variability in NPPE (Ko et al. 2020).

7 Conclusion

In this study, we examined the impact of accounting choices related to the treatment of capital on the results of FinDEA. We were motivated by the increasing number of studies using financial accounting variables to measure inputs and outputs in DEA models. Our concern was that accounting choices introduce a new source of heterogeneity into FinDEA results, but it was not clear to what extent these choices impact the results.

We used both simulated and empirical data to examine accounting choices related to the accounting measurement of capital assets in a production process. We considered two stock accounting measures of capital: the original purchase price (GPPE) and the net book value (NPPE), and one flow accounting measure: depreciation. We used GPPE as our base case, arguing that it most closely measures the physical capital used in the underlying physical production process. In contrast, depreciation and NPPE are impacted by accounting choices, such as accounting estimations (e.g., estimated useful life).

We found highly similar results when comparing the simulation and empirical results. When the FinDEA models are assumed to reflect underlying physical production processes, depreciation introduces information on the accounting estimation of an asset's useful life. NPPE further incorporates information related to an asset's age. This additional information leads to divergence in FinDEA results for different models that use the stock (GPPE and NPPE) and flow (depreciation) measures. As observed in both the simulation and the empirical test, when there is a relatively broad variation (simulation: ULC $CV = 0.25$; empirical test: ULC $CV = 0.21$) in accounting choices in measuring capital, FinDEA results using depreciation diverge from the results measured with GPPE but less than the results measured using NPPE. If GPPE is a reasonable approximation of physical capital, then the heterogeneity introduced due to different accounting choices leads to divergence between accounting and physical measures and consequently FinDEA results compared with DEA results using physical measures.

Our findings are generalizable to other accounting choices, such as inventory valuation methods, including First-in-First-out (FIFO) and Last-in-First-out (LIFO). As illustrated (Table 1), if prices are rising, FIFO assumes that the oldest inventory is sold first. This means that the costs associated with the earliest purchases are matched against revenue, leaving the newer inventory as unsold stock. This results in lower costs of goods sold, thereby enhancing profitability. However, it also means that the inventory remaining on hand is valued at the higher, more recent prices. By comparison, LIFO assumes that the most recently acquired inventory is sold first. When prices are increasing, this leads to higher costs being reported for goods sold, reducing apparent profitability. Meanwhile, inventory on hand, being valued at older, lower costs, is recorded at a lower value on the balance sheet.

The differences in valuation methods highlight the strategic trade-offs firms face. While FIFO may enhance the appearance of profitability and efficiency in FinDEA evaluations by generating higher values of the outputs (e.g. profit) relative to inputs (e.g. COGS), LIFO may provide advantages in terms of tax savings due to higher reported COGS. However, in FinDEA analysis, depending on the variable selections, this introduces complexity in how these choices affect efficiency measurement. Ultimately, the choice between FIFO and LIFO extends beyond mere accounting choices and involves strategic and operational considerations.

It is a regulatory requirement for U.S. companies using the LIFO method to disclose a LIFO reserve and outline its impact on cost of goods sold (COGS) relative to FIFO (Financial Accounting Standard Board 2015). This presents an interesting avenue for future research, which could involve: (i) conducting a FinDEA analysis for a specific industry regarding the different accounting choices among FIFO and LIFO, using the LIFO reserve; (ii) identifying firms employing LIFO and recalibrating their inventory and COGS figures using data provided in their financial statement disclosures; and (iii) reapplying the FinDEA framework to assess the differences in efficiency measurement.

We contribute to the literature in several ways. First, this study provides evidence using both simulated and empirical data on the impact of accounting choices on DEA results. We highlight accounting choices as one of the key sources of variation that lead to FinDEA results diverging from conventional DEA measures of physical productivity.

For future researchers, by enabling efficiency assessments in contexts where physical measures are unavailable or impractical, accounting data opens new avenues for using DEA to analyze productivity and performance. However, a critical insight from this research is the inherent variability introduced by accounting choices. These differences can influence the homogeneity of inputs and outputs, a foundational assumption for DEA models, and can lead to FinDEA results diverging from conventional DEA measures of physical productivity.

Future research could explore the impact of other accounting choices on FinDEA results. For example, the accounting choices for inventory (FIFO versus LIFO) or capital ownership (purchase versus lease). Future research could also explore how significant results vary or become insignificant as the structure of the FinDEA model changes. One approach to this could be reproducing the tests of existing studies and use a two-stage model to examine the statistical significance.

Second, the simulation sheds light on the extent to which accounting choices impact FinDEA outcomes. When the variation in accounting choices is relatively small (e.g., CV ≤ 0.05), FinDEA results can be interpreted as physical productivity. However, in instances of a substantial variation in accounting choices (e.g., CV ≥ 0.25), FinDEA results deviate from the base case. In such scenarios, interpreting FinDEA results as indicative of physical productivity could be misleading.

Third, our study highlights the usefulness of accounting information when researchers aim to benchmark firm performance. It emphasizes the need to carefully select accounting variables based on research objectives, as accounting choices can introduce additional performance attributes into the benchmark. In this situation, accounting information can provide valuable insights for managerial decision-making and performance evaluation.

Data availability

Data is provided within the manuscript or supplementary information files. All data are available upon request.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s11123-024-00749-9.

Author contributions Claire Cui, research design, modeling, wrote the main manuscript text, editing; Julie Harrison, research design, editing; Frederick Ng, research design; Paul Rouse, research design, modeling, editing; All authors reviewed the manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

References

Agasisti T, Munda G, Hippe R (2019) Measuring the efficiency of European education systems by combining data envelopment analysis and multiple-criteria evaluation. J Product Anal 51(2–3):105–124. https://doi.org/10.1007/s11123-019-00549-6

Ahn H, Choi S, Yun SC (2020) Financial statement comparability and the market value of cash holdings. Account Horiz 34(3):1–21. https://doi.org/10.2308/horizons-18-008

Albonico A, Kalyvitis S, Pappa E (2014) Capital maintenance and depreciation over the business cycle. J Economic Dyn Control 39:273–286. https://doi.org/10.1016/j.jedc.2013.12.008

Alhadab M, Clacher I (2018) The impact of audit quality on real and accrual earnings management around IPOs. Br Account Rev 50(4):442–461. https://doi.org/10.1016/j.bar.2017.12.003

Andor M, Hesse F (2014) The StoNED age: the departure into a new era of efficiency analysis? a Monte Carlo comparison of StoNED and the 'oldies' (SFA and DEA). J Product Anal 41(1):85–109. https://doi.org/10.1007/s11123-013-0354-y

Arya A, Fellingham J, Schroeder D (2000) Accounting information, aggregation, and discriminant analysis. Manag Sci 6(46):790–806

Baghdadi GA, Bhatti IM, Nguyen LHG, Podolski EJ (2018) Skill or effort? institutional ownership and managerial efficiency. J Bank Financ 91:19–33. https://doi.org/10.1016/j.jbankfin.2018.04.002

Baik B, Choi S, Farber DB (2020) Managerial ability and income smoothing. Account Rev 95(4):1–22. https://doi.org/10.2308/ACCR-52600

Banker RD (1993) Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. Manag Sci 39(10):1265–1273

Banker RD, Chang H (2006) The super-efficiency procedure for outlier identification, not for ranking efficient units. Eur J Operational Res 175(2):1311–1320. https://doi.org/10.1016/j.ejor.2005.06.028

Banker RD, Chang H, Natarajan R (2007) Estimating DEA technical and allocative inefficiency using aggregate cost or revenue data. J Product Anal 27(2):115–121. https://doi.org/10.1007/s11123-006-0027-1

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag Sci 30(9):1078–1092

Banker RD, Darrough MN, Huang R, Plehn-Dujowich JM (2013) The relation between CEO compensation and past performance. Account Rev 88(1):1–30. https://doi.org/10.2308/accr-50274

Banker RD, Fang S, Mehta MN (2020) Anomalous operating performance during economic slowdowns. J Manag Account Res 32(2):57–83. https://doi.org/10.2308/jmar-52547

Banker RD, Gadh VM, Gorr WL (1993) Theory and methodology a Monte Carlo comparison of two production frontier estimation methods: corrected ordinary least squares and data envelopment Analysis. Eur J Operational Res 67:332–343

Banker RD, Park H-U, Sahoo BK (2022) A statistical foundation for the measurement of managerial ability. University Library of Munich, Germany

Banker RD, Park H-U (2021) Three-stage approach to analyze managerial ability. Data Envelopment Anal J 5(1):27–54. https://doi.org/10.1561/103.00000034

Barth ME, Landsman WR, Lang M, Williams C (2012) Are IFRS-based and US GAAP-based accounting amounts comparable? J Account Econ 54(1):68–93. https://doi.org/10.1016/j.jacceco.2012.03.001

Baxter WT (1982) Lessors' depreciation and profit-an approach via deprival value. J Bus Financ Account 9(1):1–18. https://doi.org/10.1111/j.1468-5957.1982.tb00969.x

Bedford NM (1968) The foundations of accounting measurement. J Account Res 6(2):270–282

Begen MA, Ødegaard F, Sadeghi J (2023) On aggregation of technical and revenue efficiency measures. J Prod Anal, 1–16. https://doi.org/10.1007/s11123-023-00710-2

Bens DA, Monahan SJ, Steele LB (2018) The effect of aggregation of accounting information via segment reporting on accounting conservatism. Eur Account Rev 27(2):237–262. https://doi.org/10.1080/09638180.2016.1260488

Burnett BM, Cripe BM, Martin GW, McAllister BP (2012) Audit quality and the trade-off between accretive stock repurchases and accrual-based earnings management. Account Rev 87(6):1861–1884. https://doi.org/10.2308/accr-50230

Cascino S, Clatworthy MA, García Osma B, Gassen J, Imam S (2021) The usefulness of financial accounting information: evidence from the field. Account Rev 96(6):73–102. https://doi.org/10.2308/TAR-2019-1030

Charles V, Aparicio J, Zhu J (2019) The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis. Eur J Operational Res 279(3):929–940. https://doi.org/10.1016/j.ejor.2019.06.025

Charnes A, Cooper W, Rhodes E (1978) Measuring the efficiency of decision-making units. Eur J Operational Res 2:429–444

Cheng Q, Goh BW, Kim JB (2018) Internal control and operational efficiency. Contemp Account Res 35(2):1102–1139. https://doi.org/10.1111/1911-3846.12409

Ciftci M, Darrough M, Mashruwala R (2014) Value relevance of accounting information for intangible-intensive industries and the impact of scale: the US evidence. Eur Account Rev 23(2):199–226. https://doi.org/10.1080/09638180.2013.815124

Collins DW, Pungaliya RS, Vijh AM (2017) The effects of firm growth and model specification choices on tests of earnings management in quarterly settings. Account Rev 92(2):69–100. https://doi.org/10.2308/accr-51551

Cook WD, Tone K, Zhu J (2014) Data envelopment analysis: prior to choosing a model. Omega 44:1–4. https://doi.org/10.1016/j.omega.2013.09.004

Crosby N, Devaney S, Law V (2012) Rental depreciation and capital expenditure in the UK commercial real estate market, 1993–2009. J Prop Res 29(3):227–246. https://doi.org/10.1080/09599916.2012.679009

Demerjian P (2018) Calculating efficiency with financial accounting data: data envelopment analysis for accounting researchers. working paper https://ssrn.com/abstract=2995038

Demerjian P, Lev B, Lewis MF, McVay SE (2013) Managerial ability and earnings quality. Account Rev 88(2):463–498. https://doi.org/10.2308/accr-50318

Demerjian P, Lev B, McVay S (2012) Quantifying managerial ability: A new measure and validity tests. Manag Sci 58(7):1229–1248. https://doi.org/10.1287/mnsc.1110.1487

Diewert WE, Fox KJ (2023) Alternative output, input and income concepts for the production accounts. J Prod Anal. https://doi.org/10.1007/s11123-023-00701-3

Diewert WE, Mendoza MNF (1995) Data envelopment analysis: a practical alternative. UBC Department of Economics

Doumpos M, Cohen S (2014) Applying data envelopment analysis on accounting data to assess and optimize the efficiency of greek local governments. Omega 46:74–85. https://doi.org/10.1016/j.omega.2014.02.004

Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA (2001) Pitfalls and protocols in DEA. Eur J Operational Res 132(2):245–259. https://doi.org/10.1016/S0377-2217(00)00149-1

Efron B (1993) An introduction to the bootstrap. Chapman & Hall, New York

Fama EF, French KR (1997) Industry costs of equity. J Financial Econ 43(2):153–193. https://doi.org/10.1016/S0304-405X(96)00896-3

Färe R, Grosskopf S, Lovell CAK (1985) The measurement of efficiency of production. Springer, New York

Färe R, Grosskopf S, Norris M, Zhang Z (1994) Productivity growth, technical progress, and efficiency change in industrialized countries. Am Econ Rev, 84:66–83

Farrell MJ (1957) The measurement of productive efficiency. J R Stat Soc 120(3):253–290

Fields TD, Lys TZ, Vincent L (2001) Empirical research on accounting choice. J Account Econ 31(1-3):255–307

Financial Accounting Standard Board (2015) Accounting standards codification (ASC) topic 330, inventory

Golany B, Roll Y (1989) An application procedure for DEA C. Omega 17(3):237–250

Goldman E, Slezak SL (2006) An equilibrium model of incentive contracts in the presence of information manipulation. J Financial Econ 80(3):603–626. https://doi.org/10.1016/j.jfineco.2005.05.007

Harrison J, Rouse P, Armstrong J (2012) Categorical and continuous non-discretionary variables in data envelopment analysis: a comparison of two single-stage models. J Product Anal 37(3):261–276. https://doi.org/10.1007/s11123-011-0239-x

International Accounting Standards Board (2020) International Accounting Standard 16 Property, Plant and Equipment

Jones JJ (1991) Earnings management during import relief investigations. J Account Res 29(2):193–228. https://doi.org/10.2307/2491047

Jradi S, Ruggiero J (2019) Stochastic data envelopment analysis: a quantile regression approach to estimate the production frontier. Eur J Operational Res 278(2):385–393. https://doi.org/10.1016/j.ejor.2018.11.017

Kapelko M, Oude Lansink A (2017) Dynamic multi-directional inefficiency analysis of European dairy manufacturing firms. Eur J Operational Res 257(1):338–344. https://doi.org/10.1016/j.ejor.2016.08.009

Kapelko M, Oude Lansink A, Stefanou SE (2014) Assessing dynamic inefficiency of the Spanish construction sector pre- and post-financial crisis. Eur J Operational Res 237(1):349–357. https://doi.org/10.1016/j.ejor.2014.01.047

Keating AS, Zimmerman JL, Simon WE (2000) Depreciation-policy changes: tax, earnings management, and investment opportunity incentives. J Account Econ 28(3):359–389

Khezrimotlagh D, Cook WD, Zhu J (2021) Number of performance measures versus number of decision-making units in DEA. Ann Oper Res 303(1–2):529–562. https://doi.org/10.1007/s10479-019-03411-y

Kleindorfer PR, Wu DJ (2003) Integrating long- and short-term contracting via Business-to-Business Exchanges for capital-intensive industries. Manag Sci 49(11):1597–1615. https://doi.org/10.1287/mnsc.49.11.1597.20583

Ko BW, Kim J, Choi YJ, Kil KS, Lee G (2020) Enhancing the competitiveness of Korea's container shipping industry through structural improvements. Marit Policy Manag 47(1):57–72. https://doi.org/10.1080/03088839.2019.1660428

Lane J, Willett R (1997) Depreciation need not be arbitrary. Account Bus Res 27(3):179–194. https://doi.org/10.1080/00014788.1997.9729543

Lee CB, Wan J, Shi W, Li K (2014) A cross-country study of competitiveness of the shipping industry. Transp Policy 35:366–376. https://doi.org/10.1016/j.tranpol.2014.04.010

Li D, Li Y, Gong Y, Yang J (2021) Estimation of bank performance from multiple perspectives: an alternative solution to the deposit dilemma. J Product Anal 56(2–3):151–170. https://doi.org/10.1007/s11123-021-00614-z

Libby R, Rennekamp KM, Seybert N (2015) Regulation and the interdependent roles of managers, auditors, and directors in earnings management and accounting choice. Account Organ Soc 47:25–42. https://doi.org/10.1016/j.aos.2015.09.003

Ling Q, Tinkelman DP (2024) Costs and benefits of the LIFO-FIFO choice. J Corp Account Financ 35:11–20

Marquardt CA, Wiedman CI (2004) How are earnings managed? an examination of specific accruals. Contemp Account Res 21(2):461–491. https://doi.org/10.1506/G4YR-43K8-LGG2-F0XK

McCallig J, Robb A, Rohde F (2019) Establishing the representational faithfulness of financial accounting information using multiparty security, network analysis and a blockchain. Int J Account Inf Syst 33:47–58. https://doi.org/10.1016/j.accinf.2019.03.004

Oh SC, Shin J (2015) The impact of mismeasurement in performance benchmarking: a Monte Carlo comparison of SFA and DEA with different multi-period budgeting strategies. Eur J Operational Res 240(2):518–527. https://doi.org/10.1016/j.ejor.2014.07.026

Ouda HAG, Klischewski R (2019) Accounting and politicians: a theory of accounting information usefulness. J Public Budg Account Financial Manag 31(4):496–517. https://doi.org/10.1108/JPBAFM-10-2018-0113

Pilcher R, Van Der Zahn M (2010) Local governments, unexpected depreciation and financial performance adjustment. Financial Account Manag 26(3):299–324. https://doi.org/10.1111/j.1468-0408.2010.00503.x

Podinovski VV, Thanassoulis E (2007) Improving discrimination in data envelopment analysis: some practical suggestions. J Product Anal 28(1–2):117–126. https://doi.org/10.1007/s11123-007-0042-x

Portela MCAS, Thanassoulis E (2014) Economic efficiency when prices are not fixed: disentangling quantity and price efficiency. Omega 47:36–44. https://doi.org/10.1016/j.omega.2014.03.005

Ren XT, Yang GL (2023) Eliminating congestion in China's paper-making and paper products industry: from both the perspective of increasing and decreasing inputs. J Prod Anal. https://doi.org/10.1007/s11123-023-00675-2

Rodríguez-Pérez G, Slof J, Solà M, Torrent M, Vilardell I (2011) Assessing the impact of fair-value accounting on financial statement analysis: a data envelopment analysis approach. Abacus 47(1):61–84. https://doi.org/10.1111/j.1467-6281.2011.00331.x

Rouse P, Tripe D (2016) Allocative and technical efficiency of New Zealand banks. Meditari Account Res 24(4):574–587. https://doi.org/10.1108/MEDAR-02-2016-0035

Ruggiero J (1999) Efficiency estimation and error decomposition in the stochastic frontier model: a Monte Carlo analysis. Eur J Operational Res 115(3):555–563

Ruggiero J (2004) Data envelopment analysis with stochastic data. J Operational Res Soc 55(9):1008–1012. https://doi.org/10.1057/palgrave.jors.2601779

Ruggiero J (2007) A comparison of DEA and the stochastic frontier model using panel data. Int Trans Operational Res 14(3):259–266. https://doi.org/10.1111/j.1475-3995.2007.00585.x

Schwab CM, Stomberg B, Williams BM (2022) Effective tax planning. Account Rev 97(1):413–437. https://doi.org/10.2308/TAR-2019-0020

Shephard RW (1970) Theory of cost and production functions. Princeton University Press, New Jersey

Simper R, Hall MJB, Liu W, Bin, Zelenyuk V, Zhou Z (2017) How relevant is the choice of risk management control variable to non-parametric bank profit efficiency analysis? the case of South Korean banks. Ann Oper Res 250(1):105–127. https://doi.org/10.1007/s10479-015-1946-x

Smith P (1997) Model misspecification in data envelopment analysis. Ann Oper Res 73:233–252. https://doi.org/10.1023/a:1018981212364

Taborda D, Sousa J (2023) The accrual accounting principle and its implications for portuguese tax courts decisions. Account Econ Law Conviv 13(4):539–562. https://doi.org/10.1515/ael-2019-0030

Teoh SH, Welch I, Wong TJ (1998) Earnings management and the long-run market performance of initial public offerings. J Financ 53(6):1935–1974. https://doi.org/10.1111/0022-1082.00079

Thomas AL (1969) Studies in accounting research No.3. American Accounting Association, Chicago

Thomas AL (1974) Studies in accounting research No.9. American Accounting Association, Chicago

Vander Bauwhede H, Willekens M, Gaeremynck A (2003) Audit firm size, public ownership, and firms' discretionary accruals management. Int J Account 38(1):1–22. https://doi.org/10.1016/S0020-7063(03)00004-9

Wan Z, Zhu M, Chen S, Sperling D (2016) Pollution: three steps to a green shipping industry. Nat News, 530:275–277

Weygandt JJ, Kimmel PD, Kieso DE (2015) Financial accounting: IFRS edition. Wiley, Singapore

Wilhelmsson M, Roos H (2024) Age-related depreciation of older properties. J Prop Res 41(1):44–70. https://doi.org/10.1080/09599916.2023.2296442

Young S (1999) Systematic measurement error in the estimation of discretionary accruals: an evaluation of alternative modeling procedures. J Bus Financ Account 26(7):833–862

Yu W, Ramanathan R, Nath P (2014) The impacts of marketing and operations capabilities on financial performance in the UK retail sector: a resource-based perspective. Ind Mark Manag 43(1):25–31. https://doi.org/10.1016/j.indmarman.2013.07.014

Zelenyuk V (2023) Productivity analysis: roots, foundations, trends and perspectives. J Product Anal 60(3):229–247. https://doi.org/10.1007/s11123-023-00692-1