A Conceptual Foundation for Financial DEA

Claire Mu-Rong Cui

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Accounting, the University of Auckland, 2021.

Abstract

This study aims to provide a conceptual foundation for Financial Data Envelopment Analysis (DEA), which refers to DEA models that exclusively use accounting data from external financial reports to measure firm performance. Researchers have argued that using accounting measures in DEA models without appropriate methodological rationalisation might distort performance evaluations (Färe et al., 1985; Färe et al., 2017). Researchers have also questioned the kind of efficiency measures being calculated by DEA when the inputs and outputs are accounting measures since they cover both quantities and prices (Banker et al., 2007; Cross & Färe, 2008; Portela, 2014; Zelenyuk, 2020). This study is motivated by the growth in Financial DEA research, where DEA model are potentially being applied without underpinning by an articulated conceptual foundation

A two-phase analytical approach is used to examine the uses of Financial DEA and develop a conceptual foundation to assist the design and interpretation of Financial DEA. Phase I provides a typology of Financial DEA literature and reviews the methodological issues in the application of Financial DEA. Phase II quantifies selective methodological issues with empirical illustrations, using Monte Carlo simulations and analysis of archival data.

The key findings and contributions are three-fold. First, the typology identifies 12 dimensional constructs of firm performance and nine indicators. This contributes a conceptual framework at the construct level, which provides an overview of the scope of Financial DEA and can be used to position the application of Financial DEA and apprise Financial DEA practice. Second, the methodological issues are formed into a four quadrant framework of measurement models with various measurement errors in Financial DEA. This contributes a conceptual framework at the modelling level, which can be used to highlight potential methodological pitfalls in Financial DEA. Third, the empirical tests demonstrate the quantitative magnitudes of selective methodological issues on Financial DEA results, exploring how research contexts affect results to varying degrees. This contributes a conceptual framework at the factor level, which can guide accounting variable selection in Financial DEA practice so that the measurement errors can be reduced. These three frameworks form a conceptual foundation for future studies of Financial DEA.

In memory of my grandmother, Suzhen Zhang.

To my parents, partner, and family.

Acknowledgements

It feels like yesterday I enrolled on my PhD, and now I am standing at the last step before submission. Everything has gone by so fast before my eyes, just like reading a book and then realising the day has gone. I thought my PhD was going to be a journey. However, now I realise it is the first step of a lifelong pursuit.

I would like to express my greatest gratitude to my supervisors and advisor, Associate Professor Julie Harrison, Dr Frederick Ng, and Professor Paul Rouse. This work is impossible without their guidance, patience, and caring. I have learned so much as a student, thanks to their knowledge and wisdom. Their passion for research is compelling, and I hope to follow in their footsteps in the future.

I appreciate the research enthusiasm and experience shared by those from the Accounting and Finance Department and the Business School: Professor Julie Cassidy, Professor Kevin Dow, Associate Professor Jeff Casterella, Dr Jonathan Albright, Dr Sharlene Biswas, Dr Lily Chen, Dr Caro Crowhurst, Dr Maryam Hasannasab, Dr Michelle Li, Dr Angela Liew and Dr Ramona Zharfpeykan. I am grateful to Professor Norman Wong, the Head of Department, for his kind encouragement throughout my student life. I would also thank the department manager Herena for her constant caring.

I would like to thank the many other parties who have contributed to this work. The funding from PhD conference funding from the Business School, and the 2021 Business School Student Support Award. The constructive comments and precious feedback from the Institute for Operations Research and the Management Sciences (INFORMS), Chartered Institute of Management Accountants (CIMA), New Zealand Management Accounting (NZMA), and Auckland Region Accounting (ARA) conferences.

I would like to give special thanks to my friends: Lina, Isabella, Freya, Wanyi, Jun, Shandong, Diamond, Lucy, Jonnie, Eva, and Nancy, for their priceless friendship and all the tears and the laughter we shared.

I would like to thank my family with my deepest love: my mum, Qingfeng; my dad, Yanfeng; my grandparents, Junwei, Suzhen, Shuzhen and Xiuzhi; and my partner, Zack. Thank them for letting me be myself, respecting my curiosity about the world, and being with me.

Most importantly, I would like to thank my grandmother, Suzhen, who passed away during the last few months of my thesis write up. It is a pity I could not fly back home to say goodbye to her for the last time. However, thank you, grandma, for being a role model for me, for looking after me during my childhood, and for encouraging me to pursue my research dream. Hopefully, you will be pleased to hear that your granddaughter has become a doctor now and fulfilled her promise to you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACT: current assets

AE: allocative efficiency

AGE: age of property plant and equipment

AT: total assets

CAPEX: capital expenditures

CEQ: common equity

CEQ-PO: common equity, translated to positive values

COGS: cost of goods sold expense

CRS: constant returns to scale

DEA: data envelopment analysis

DMU: decision-making unit

DP: depreciation expense

EBIT: earnings before interest and taxes

EBITDA: earnings before interest, taxes, depreciation and amortisation

EPS: earnings per share

Financial DEA: DEA empirical research that only uses financial statement data as inputs and outputs

GDWL: goodwill

GPPE: gross property, plant and equipment

IAS: International Accounting Standard

IASB: International Accounting Standards Board

IFRS: International Financial Reporting Standard

INTAN: intangible assets

LT: total liabilities

MCDM: multiple criteria decision making

MKVALT: market value

MRC: capitalised operating lease expenses

Net R&D: net research and development expense

NI: net income

NI-PO: net income, translated to positive values

NOPAT: net operating profit after tax

NPPE: net property, plant and equipment

OE: overall efficiency

SALE: sales revenue

SG&A: selling, general, and administrative expenses

TE: technical efficiency

UL: useful life of property plant and equipment

VRS: variable returns to scale

XOPR: operating expenses

XRD: capitalised research and development expense

XSGA: selling, general and administrative expenses

Chapter 1: Introduction

1.1. Chapter Introduction

Data Envelopment Analysis (DEA) is a performance measurement method developed based on economic productivity theories and operations research (Banker et al., 1984; Charnes et al., 1978; Farrell, 1957). DEA incorporates multiple measures of inputs and outputs to measure the performance of decision-making units (DMUs) such as firms and organisations without requiring knowledge of the production function. DEA research is widely used in methodological studies which develop theoretical concepts and examine methodological issues. DEA is also used in empirical studies, in which DEA is applied in a wide variety of settings such as banking, healthcare, agriculture, transportation, and education (Cook & Seiford, 2009; Liu et al., 2012).

In recent years, an emerging stream of DEA empirical research has only incorporated monetary measures as inputs and outputs, termed "Financial DEA" in this study.[1] Financial DEA has been used mainly to measure firm performance using accounting data. As a result, Financial DEA is a measurement model located at the nexus of the DEA algorithm, accounting information, and firm performance measurement.

Financial DEA is different from conventional DEA applications, which incorporate physical measurements to measure productive efficiency, as defined by Farrell (1957). Researchers have argued that using accounting measures in DEA models without appropriate methodological rationalisation might distort performance evaluations (Färe et al., 1985; Färe et al., 2017; Portela, 2014; Zelenyuk, 2020). Such distortions may arise where Financial DEA applications are not conceptually underpinned by production processes or benchmarking operations, thereby possibly rendering the results meaningless (Cook et al., 2014). Harrison and Rouse (2016) proposed that alternative specifications of Financial DEA might capture different perspectives of performance. Therefore, researchers need to be cognizant of the conceptual foundation of Financial DEA.

[1] In this study, the word "monetary" and "financial" are used interchangeably for Financial DEA variables.

This study seeks to provide a conceptual foundation of Financial DEA comprising a typology of performance conceptualisations and a structured list of methodological issues arising from Financial DEA application. This is developed by following a two-phase research plan. In Phase I, the conceptualisation phase will examine the development of Financial DEA using a literature survey. The output will be a typology of Financial DEA studies at the construct level. Methodological issues arising when using Financial DEA will also be examined, with the output being identification of a range of issues in the Financial DEA application process. Based on Phase I, Phase II examines selective methodological issues covered by Financial DEA literature. The output will be an assessment of the nature and magnitude of measurement errors under various research conditions in the application of Financial DEA.

1.2. Background

This section provides a background to the study and provides an overview of the current Financial DEA research. Financial DEA is defined in this study as a model that utilises the DEA technique and accounting information exclusively (rather than physical measures) to measure performance.

Utilising the mechanics of the DEA technique and the large quantity of public financial data accessible from electronic databases, Financial DEA research started to emerge in the 1990s. As illustrated in Figure 1 - 1, 248 Financial DEA models (210 studies) have been identified, with the majority published in the last fifteen years.

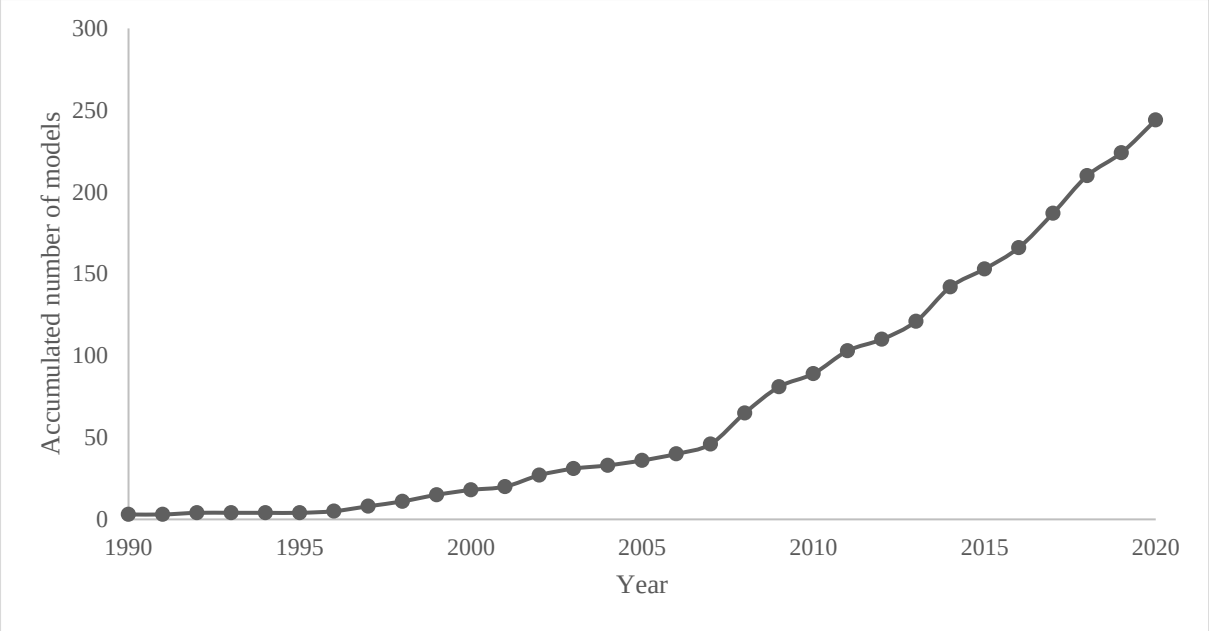


Figure 1 - 1 Emerging Trend in Financial DEA Studies

A seminal study by Smith (1990) is the first Financial DEA study that attempted to use Financial DEA as a complementary tool for financial ratio analysis. The study has expanded the toolkit of financial statement analysis to utilise the DEA technique. In more recent years, Demerjian et al. (2012) and Demerjian et al. (2013) used Financial DEA to measure firm performance and a second-stage regression to calculate managerial ability. The DEA models in these two studies covered a wide range of accounting variables and applied them to multiple industries with large panel data over long time series. These studies have drawn financial accounting researchers' attention to Financial DEA, and they have been highly cited over the years.[2]

The scope of Financial DEA covers various dimensions of firm performance. For example, to measure productive efficiency, accounting measures are used to closely proxy physical production elements in a specific physical production process as in conventional DEA (Wu et al., 2016). Operational efficiency measures the productive efficiency of multiple aggregated production processes, and the accounting measures proxy the physical production remotely (Aparicio & Kapelko, 2018). Funding efficiency provides capital support of firm performance, and the accounting measures relate to capital structures (Oberholzer, 2014b;

[2] Demerjian et al. (2012) has been cited for 974 times, and Demerjian et al. (2013) has been cited for 751 times as at 25th May 2021.

Smith, 1990). Profitability efficiency measures the efficiency of generating revenues from expenses, where the accounting variables do not seem to relate to specific physical production processes closely. Rather, firm performance is benchmarked to minimise expenses while maximising revenues (Kaffash et al., 2018). Financial efficiency minimises financial resources while maximising the financial outputs (Demerjian et al., 2012; Demerjian et al., 2013). The accounting variables relate to the production process remotely, indicating a benchmarking rather than production efficiency approach (Baghdadi et al., 2018). In this sense, Financial DEA is used as a benchmarking tool similar to financial ratio analysis (Feroz et al., 2001; Feroz et al., 2003).

However, there are limitations of Financial DEA. For example, the accounting variable selection needs a theoretical grounding (Banker et al., 2021; Bowlin, 1999), such as productivity or performance benchmarks. Also, the accounting variables may expose results to factors such as inflation and variations in accounting practices (Banker et al., 2021; Smith, 1990). Further, DMUs may be impacted by specific operating environments and firm- and industry-specific variables and so be less comparable in DEA (Demerjian, 2018; Rahman et al., 2019; Smith, 1990). Also, the accounting variables may not appear in sufficient disaggregated forms to measure firm performance (Bowlin, 1999; Liu et al., 2018; Smith, 1990). Last, on the DEA technique side, researchers have questioned the nature of efficiencies measured by DEA when accounting variables are used (Färe et al., 2017; Portela, 2014).

1.3. Motivation

The motivation for this study is twofold. First, there has been a rapidly increasing number of Financial DEA studies since Smith (1990). However, there are fragmented views of Financial DEA since the applications lack a comprehensive overview of how Financial DEA studies have been used. This study structures the Financial DEA literature by providing a typology, which categorises the dimensional constructs of firm performance measured by Financial DEA. It also categorises the indicators and accounting variables used in Financial DEA models. It examines the Financial DEA research design and modelling and provides the basis for developing a conceptual framework.

Second, Financial DEA is a complementary tool to traditional statistical analysis in accounting research (Bowlin, 1999; Demerjian et al., 2012). Demerjian et al. (2012) suggested using Financial DEA as a complement to traditional regression methods. Bowlin (1999) also suggested that Financial DEA can be used to complement financial ratio analysis. However, several difficulties have hindered Financial DEA from wider applications. Accounting data sources are often limited to public data at a highly aggregated level. These aggregated accounting data may not measure firm performance with sufficient detail (Bowlin, 1999; Liu et al., 2018; Smith, 1990). Researchers have questioned the nature of efficiencies generated by price-based aggregated accounting data in DEA (Färe et al., 2017; Portela, 2014). This research responds to the call for facilitating the application of Financial DEA (Harrison & Rouse, 2016; Joo et al., 2011). The study reviews the potential methodological issues in the application of Financial DEA.

This study responds to calls for research on the conceptual issues of Financial DEA (Banker et al., 2021; Bowlin, 1999; Harrison & Rouse, 2016). For the past three decades, there have been increasing numbers of empirical Financial DEA studies (Demerjian, 2018; Hasan, 2020). Nevertheless, few studies have discussed the conceptual issues of Financial DEA. For example, Bowlin (1999) criticised the lack of proven theoretical construct for the variables used in the Financial DEA models. Banker et al. (2021) commented that the factors which drive productivity need further investigation. For example, Banker et al. (2021) questioned whether technology spillover partially explains the increase in productivity. This study aims to provide a comprehensive conceptual foundation of Financial DEA and guide its future application.

1.4. Research Questions

The overarching goal of this research is to develop a conceptual foundation for Financial DEA. The study will examine the following three research questions to achieve this goal:

> ***RQ1.*** *What dimensions of firm performance do Financial DEA models measure, and how have they been used?*

In the literature, Financial DEA models have been used to measure various dimensions of firm performance. For example, the productive efficiency of a specific production process was measured with accounting variables used as proxies of physical productive elements

(Giokas, 2008; Kweh et al., 2018). Funding efficiency was used to measure the efficiency of the capital structure supporting firm performance (Oberholzer, 2014b; Smith, 1990). Financial efficiency was measured to minimise financial resources while maximising the financial outputs (Demerjian et al., 2012; Demerjian et al., 2013).

This study surveys the Financial DEA literature to develop a typology of Financial DEA that can inform the scope of Financial DEA to answer this question. The study examines the variety of dimensional constructs of Financial DEA models measured in the literature. The output of this research question is a comprehensive typology, including the categories of dimensional constructs, indicators, and accounting variables used in the empirical Financial DEA literature to date. The typology also delivers a conceptual framework at the construct level, covering the scope of dimensional constructs of firm performance that the literature has measured.

RQ2. *What are the methodological issues when applying Financial DEA?*

There are various methodological issues to be considered during the application of Financial DEA. For example, the homogeneity assumption requires DMUs to be comparable with regard to the activities, outputs, and environment (Dyson et al., 2001). Financial DEA applications are exposed to heterogeneity due to: the operating environments, firm- and industry-specific variables (Demerjian, 2018; Rahman et al., 2019; Smith, 1990), accounting practices, and accounting regulations (Banker et al., 2021; Smith, 1990). Other assumptions include sample size versus discriminatory power (Demerjian, 2018), non-negative values (Bowlin, 1999; Feroz et al., 2003), and variable selection conceptual basis (Banker et al., 2021; Bowlin, 1999). Also, researchers have raised concerns about the nature of the efficiency measured by the Financial DEA when price-based aggregated accounting variables are used in DEA (Färe et al., 2017; Portela, 2014).

This study will review measurement models and Financial DEA models to examine the methodological issues arising in the Financial DEA application process to answer this research question. The output of this research question is a conceptual framework at the modelling level, organising the potential methodological issues and measurement errors in the Financial DEA application process.

RQ3. *What are the empirical impacts of methodological choices on the results of Financial DEA?*

This study will examine the quantitative impact of selective measurement issues identified from the literature on Financial DEA with empirical data sets. Three sets of empirical studies will be conducted. The first test will examine the impact of input and output prices on the Financial DEA results. In conventional DEA, physical measures calculate technical efficiency (Farrell, 1957). However, researchers have raised concerns about the impact of price-based aggregated accounting data used in Financial DEA (Portela, 2014). There have been many studies conceptually discussing the impact of price factors on the results of DEA (Färe et al., 1990; Färe et al., 2017). However, few studies provide empirical evidence about the impact of price factors on the results of Financial DEA.

The second test will examine the impact of alternative stock and flow forms of accounting variables on the Financial DEA results. Researchers suggest the need to investigate the impact of accounting practices on the results of Financial DEA (Banker et al., 2021; Smith, 1990).

The third test will examine the impact of alternative accounting variables and alternative dimensions of firm performance on the Financial DEA results across various industry settings. The researchers have suggested that in Financial DEA, the DMUs could be less comparable due to the heterogeneous operating environment, firm, and industry features (Demerjian, 2018; Rahman et al., 2019; Smith, 1990). However, there are relatively few Financial DEA studies that compare the heterogeneity across industries.

The output of this research question is a conceptual framework at the factor level in the context of the business environment. The framework will also provide selective empirical evidence of methodological issues to guide future Financial DEA applications.

1.5. Expected Contributions

The contributions of this research are threefold. First, the study will structure the Financial DEA literature and provide a typology of Financial DEA applications. This typology provides a summary of the development of Financial DEA studies to-date. The typology forms a conceptual framework at the construct level, which describes the scope of Financial

DEA research. It covers the diversity of dimensional constructs of firm performance measured by Financial DEA and the pattern of indicators and accounting variables used.

Second, this study will examine methodological issues of Financial DEA applications. Various Financial DEA issues may hinder Financial DEA from being interpreted meaningfully. This study will explore issues using measurement models to understand various types of measurement errors. The methodological issues discovered are organised into a conceptual framework at the modelling level, which assists in locating the source of measurement errors in the modelling process.

Third, this study will provide selective empirical evidence of the impact of various methodological choices on Financial DEA applications. The selective empirical evidence provides a conceptual framework at the factor level, facilitating the application of Financial DEA in different business contexts.

Overall, this study provides a comprehensive conceptual foundation at the construct, modelling, and factor levels. The conceptual foundation can be used to guide the adoption of Financial DEA as a firm performance measurement tool.

1.6. Chapter Summary

In summary, this chapter introduces the research. This study aims to develop a conceptual foundation of Financial DEA. The number of Financial DEA studies has been rapidly increasing in recent years; however, the conceptual foundation is unclear. It is important to understand the conceptual foundation of Financial DEA to test theories meaningfully. This study reviews the scope of Financial DEA and explores the methodological issues in Financial DEA conceptually and empirically. This study provides conceptual frameworks at the construct, modelling, and factor level to provide a comprehensive conceptual foundation of Financial DEA.

The study is structured as follows. Chapter 2 reviews the literature on DEA, accounting information, firm performance, and measurement models. These themes form the Financial DEA domain and provide the prerequisite knowledge to develop a conceptual foundation of Financial DEA applications. The literature review provides an overview of Financial DEA

and motivates the need for a literature survey of Financial DEA that describes the scope of its development to date. The literature survey is discussed in Chapter 4.

Chapter 3 presents the research approach for the study. It identifies the paradigm of this study. This chapter details the methods used in the two phases of the research.

Chapter 4 details the findings of the conceptualisation phase (Phase I). It covers a typology of constructs, indicators, and variables discovered in the Financial DEA literature survey. It also reports the methodological issues relevant to the application of Financial DEA, based on which a range of propositions are developed to be tested in Chapter 5.

Chapter 5 covers the findings from the empirical examination phase (Phase II). It covers three sets of empirical tests selected from the methodological issues identified in the Financial DEA literature in Chapter 4. The three sets of tests examine the impact of (a) price factors, (b) alternative forms of stock and flow accounting variables, and (c) alternative accounting variables in various business settings.

Chapter 6 discusses and synthesises the outputs of the three research questions in Chapters 4 and 5. This chapter synthesises the frameworks at the construct, modelling, and factor levels to provide a comprehensive conceptual foundation to facilitate the application of Financial DEA.

Chapter 7 concludes the study. It summarises the key findings from the three research questions and their contributions. It also discusses the limitations of the research and suggests directions for future research to examine the conceptual foundation of Financial DEA further.

Chapter 2: Literature Review

2.1. Chapter Introduction

This chapter reviews the relevant DEA and accounting literature to develop a conceptual framework of Financial DEA. The themes of the literature are selected to motivate the overarching research goal. This chapter focuses on the literature needed to support the examination of Financial DEA being the nexus of DEA, accounting information, firm performance, and measurement models. A synthesis of findings from these fields provides the fundamental underpinning of Financial DEA. Section 2.2 introduces the fundamentals of DEA and provides the key definitions, the history, and the key features of the algorithm. Section 2.3 reviews the nature of the accounting information related to the data processed by Financial DEA. The nature of the accounting information distinguishes Financial DEA from conventional DEA, which processes physical data. Understanding the nature of the accounting information provides insights into conceptual and methodological issues arising from the use of accounting information in Financial DEA. Section 2.4 reviews firm performance to provide a context to this study. Section 2.5 reviews measurement models: reflective models and formative models and investigates how these models can be adapted to Financial DEA research. These two types of models exist at different abstraction levels in the Financial DEA research process. Section 2.6 summarises the chapter.

2.2. Data Envelopment Analysis

2.2.1. Overview of DEA

DEA is an optimisation tool to measure the relative performance (relative efficiencies) of organisations or parts of organisations (Charnes et al., 1978; Debreu, 1951; Farrell, 1957; Koopmans, 1951). The organisations being measured, including for-profit and not-for-profit types, are termed decision-making units (DMUs) (Charnes et al., 1978).

DEA is mathematically based on linear programming, where weights are unknown variables applied to inputs used and outputs produced by a DMU in its production process. By solving the linear programme to obtain optimal weights, each DMU is assigned an optimal efficiency score relative to all other DMUs in the comparison group. The DMUs with the highest

efficiency scores (100% efficiency) can be used to construct a piece-wise frontier from this process, which is considered the efficient frontier or frontier of best performance. Within the efficient frontier are the relatively inefficient DMUs. The distance from inefficient DMUs to surrounding efficient DMUs indicates the strategies for performance improvement. To be a better performer, the inefficient DMUs need to learn from their peers who are the efficient DMUs with similar resource allocations and strategic goals (Coelli et al., 2005; Cooper et al., 2011; Rouse et al., 2010)

DEA has several advantages as a tool for performance measurement. First, DEA mathematically allows for free weights. This feature enables DEA to calculate the efficiency scores by accommodating each DMUs' strengths while prohibiting opportunistic behaviours. For example, opportunistic firms may focus on the attributes with heavy weights if the weights are predetermined. Compared with other performance measurement metrics, especially those with fixed weights, DEA, as a non-parametric tool, accommodates different production approaches. Second, DEA mathematically allows for multiple inputs and outputs in a model; therefore, it can calculate multi-dimensional firm performance compared with other quantitative firm performance measurement tools, such as financial ratios, which can be limited to a single numerator and denominator, DEA can simultaneously incorporate multiple dimensions. Third, after incorporating multiple dimensions, DEA generates a single composite efficiency score for benchmarking. Fourth, DEA is a non-parametric model, which does not require pre-specification of the underlying production function. Compared with parametric functions, such as Corrected Ordinary Least Squares (COLS), DEA has the advantage that it does not require specification of the production function. Fifth, DEA is a frontier analysis tool. Compared with regressions, which focus on the central tendency, DEA identifies the best performers as peers from which the less efficient DMUs can learn and improve.

Because of these advantages, DEA has been widely applied as a performance measurement tool in economics, operations research, and accounting fields and various industries, such as the banking industry (Raith et al., 2019), healthcare (Färe et al., 2007), agriculture (Färe et al., 2006), and education (Avilés-Sacoto et al., 2014).

2.2.2. History of DEA

DEA was first introduced by Charnes et al. (1978), based on the seminal work by Farrell (1957). Farrell (1957) used partial productivities such as labour productivity and capital productivity to develop a comprehensive productivity model by combining multiple inputs.

Farrell (1957), based on earlier work on productivity (Debreu, 1951; Koopmans, 1951), developed a model to find the Pareto optimal or efficient production frontier, which is the point where no one can be better off without making others worse off (Cooper et al., 2011).

Farrell (1957) distinguished the term technical efficiency (TE) from overall efficiency (OE) and allocative efficiency (AE). First, TE is concerned with maximising the physical outputs or minimising physical inputs. Second, AE considers how to optimise efficiency when taking price factors into account. Third, OE is the product of TE and AE.

Following Farrell (1957), physical measures are used to measure TE, which is defined as "success in producing as large as possible an output from a given set of inputs" (Farrell, 1957, p. 254). By comparison, AE includes price factors, and as Farrell pointed out, "price (allocative) efficiency measures the extent of a firm's adaptation to a particular set of prices" (Farrell, 1957, p. 261). The OE "is equal to the product of the technical efficiency and price (allocative) efficiencies" (Farrell, 1957, p. 255). Coelli et al. (2005, p. 183) further distinguished the differences between TE and AE; they explained that "if price data are available and a behavioural objective, such as cost minimisation or revenue or profit maximisation, is appropriate, then it is possible to measure allocative efficiencies as well as technical efficiencies".

Farrell (1957) highlighted that the calculation of AE requires accurate price information. The differences between the three types of efficiencies (TE, OE, and AE) are vague in the absence of price information. Unstable AE can lead to problems with interpretation (Farrell, 1957). Farrell (1957) also pointed out that AE is hard to interpret as high AE is not necessarily desirable. For example, during a firm's expansion, low AE may be needed to increase output production.

Figure 2 - 1 is used to illustrate how TE, AE, and OE differ. The figure is from Coelli et al. (2005), which took the data of five firms (1-5 in Figure 2 - 1) from a two-input, one-output,

input-oriented constant returns to scale (CRS) DEA model.[3] The two inputs are noted as x_1 and x_2 and the output is noted as q in the figure. All the firms face the same prices, which are 1 and 3 for input x_1 and x_2 respectively. The prices form an isocost line with a slope of -1/3, where firms are allocatively efficient. The firms on the frontier are technically efficient. Among the five firms, only firm 5 is overall efficient and technically efficient, without allocative inefficiency.

The calculation of TE, AE, and OE can be illustrated by firm 3. The TE is measured as the ratio of the distance from 0 to the point $3'$ divided by the distance from 0 to 3, and the result of TE is 0.833. The AE is calculated as the ratio of the distance from 0 to $3''$ divided by the distance from 0 to $3'$, and the result of AE is 0.9. The OE is the ratio of the distance from 0 to $3''$ divided by the distance from 0 to 3, and the result of OE is 0.75. The values of OE, AE, and TE follow $OE = AE \times TE$ $(0.75 = 0.9 \times 0.833)$.

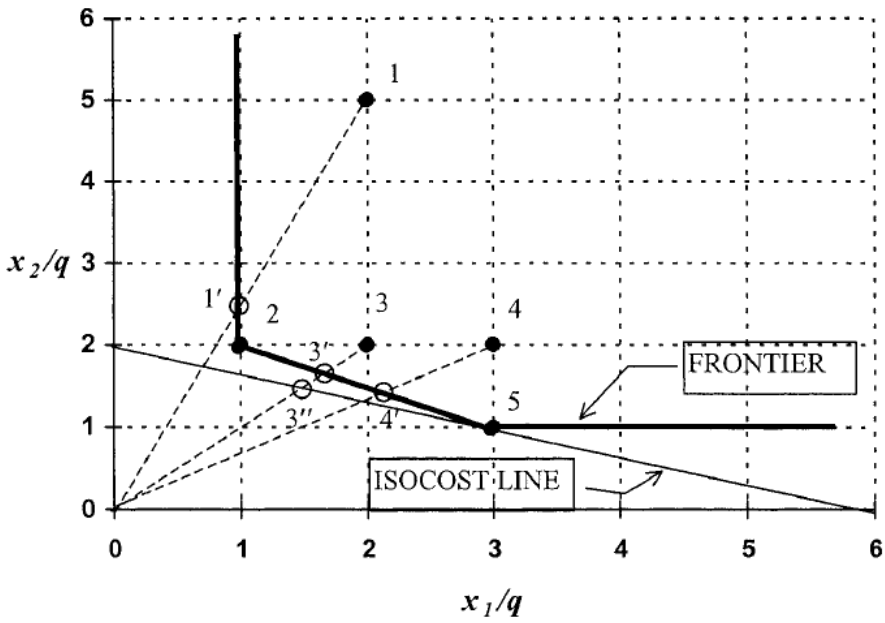


Figure 2 - 1 Technical Efficiency, Allocative Efficiency, and Overall Efficiency from Coelli et al. (2005)

[3] There are two types of DEA orientation, input-orientation, and output-orientation. Input-orientation focuses on the efficiency of consuming inputs while holding the outputs constant. On the other hand, output-orientation focuses on the efficient generation of outputs while inputs are fixed (Charnes et al., 1978). There are two types of returns to scale: the constant returns to scale (CRS) and the variable returns to scale (VRS) model (Charnes et al., 1978, Banker et al., 1984). More details are in section 2.2.2.1 and section 2.2.2.3.

In the context of Financial DEA, the distinction among the three efficiencies (OE, TE, AE) tends to be vague. On the one hand, Financial DEA could measure OE since accounting information is the product of physical quantities and prices.[4] On the other hand, TE could be measured if accounting measures are treated as proxies of physical quantities (Harrison & Rouse, 2016).

2.2.2.1. ***Charnes, Cooper and Rhodes (1978): CCR model***

Charnes et al. (1978), based on the work of Farrell (1957), extended the single output model presented in Farrell (1957) to include multiple inputs and outputs and developed the DEA model in its present form.

Charnes et al. (1978) introduced DEA in the ratio form (input orientation), as the ratio of outputs to inputs (Cooper et al., 2011). The ratio formed is stated below[5]:

$$\max h_0(\mathrm{u},\mathrm{v}) = \frac{\sum_r u_r y_{ro}}{\sum_i v_i x_{io}} \qquad (2-1)$$

subject to

$$\frac{\sum_r u_r y_{rj}}{\sum_i v_i x_{ij}} \leq 1 \; for \; j = 1, \dots, n,$$

$$u_r, v_i \geq 0 \; for \; all \; i \; and \; r.$$

Assuming there are in total *n* DMUs to be evaluated, each DMU consumes *m* types of inputs to produce *s* types of outputs. $\mathrm{DMU_j}$ consumes amount x_{ij} of input i and produces amount of y_{rj} of output r. The *i*th type of input of $\mathrm{DMU_j}$ is denoted as $x_{ij} \geq 0$ for *m* types of inputs. Similarly, the *r*th type of output of $\mathrm{DMU_j}$ is denoted as y_{rj}, $y_{rj} \geq 0$ for *s* types of outputs (Cooper et al., 2011).

The ratio form yields an infinite number of solutions. The transformation of the ratio form for linear fractional programming selects a solution (u,v) for which $\sum_{i=1}^{m} v_i x_{io} = 1$. The ratio

[4] Prices may not be current prices.
[5] Adapted from Charnes et al. (1978) and Cooper et al. (2011). The equation 2 - 1, 2 - 2 and 2 - 3 are input-orientation.

form of DEA is changed to a linear programming problem in the multiplier form (input orientation)[6]:

$$\max z = \sum_{r=1}^{s} \mu_r y_{ro} \qquad (2-2)$$

subject to

$$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0$$

$$\sum_{i=1}^{m} v_i x_{io} = 1$$

$$u_r, v_i \geq 0$$

The change of variables from (u,v) to (μ, ν) is a result of the "Charnes - Cooper" transformation (Cooper et al., 2011).

After taking the dual of the equation, DEA is transformed to the envelopment form (input orientation), as stated below[7]:

$$\theta^* = min\theta \qquad (2-3)$$

subject to

$$\sum_{j=1}^{n} x_{ij}\lambda_j \leq \theta x_{io} \qquad i = 1,2,\dots,m;$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{ro} \qquad r = 1,2,\dots,s;$$

$$\lambda_j \geq 0 \qquad \mathrm{j} = 1,2,\dots,\mathrm{n}.$$

[6] Adapted from Cooper et al. (2011).
[7] Adapted from Cooper et al. (2011).

In the envelopment form, the λ is a vector of intensity variables denoting the linear combination of DMUs. The objective function θ is a radial contraction factor that can be applied to DMU_o's inputs (Paradi et al., 2018).

This dual linear programme in DEA is based on the production function in economics. Duality theory transforms the multiplier form to the envelopment form, and therefore, they are a primal-dual pair of the linear programme. Thus, the optimal solution and efficiency scores are the same. Duality theory is explained in more detail below.

2.2.2.2. ***<u>Duality theorems</u>***

Duality theory derives from the Lagrangian dual problem in mathematical programming (Shephard, 1953; Shephard, 1970). Conventionally, the primal space is a quantity space, and the dual space is a price space (Färe et al., 2017).

In DEA and production theory, duality theory can be explained through the relationship between the multiplier form of DEA models (dual function) and the envelopment form of DEA models (primal function). In DEA, the primal (quantity) space is expressed in the envelopment form. Equation (2 - 4) repeats equation (2 - 3) and provides the CRS input-orientation model in the envelopment form:

$$\theta^* = min\theta \qquad (2-4)$$

subject to

$$\sum_{j=1}^{n} x_{ij}\lambda_j \leq \theta x_{io} \qquad i = 1,2,\dots,m;$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{ro} \qquad r = 1,2,\dots,s;$$

$$\lambda_j \geq 0 \qquad \mathrm{j} = 1,2,\dots,\mathrm{n}.$$

In the envelopment form, the focus is the quantities of inputs and outputs, and the objective function θ is a scalar. The θ obtained is the efficiency score for the DMU under evaluation.

The λs are the combinations of the efficient DMUs that form the reference set, and the constraint sets the target for a DMU to be technically efficient.[8] The DMUs are efficient in the sense that no other DMUs can offer a reduction in one of its input levels without either a consequent rise in some other input level or a reduction in the level of at least one of the outputs (Thanassoulis, 1997). The envelopment model can infer prices via the shadow prices of the constraints.

The dual (price) space is expressed in the multiplier form DEA model Equation (2 - 5) repeats (2 - 2) provides the CRS input-orientation model in the multiplier form:

$$\max z = \sum_{r=1}^{s} \mu_r y_{ro} \qquad (2-5)$$

subject to

$$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0$$

$$\sum_{i=1}^{m} v_i x_{io} = 1$$

$$u_r, v_i \geq 0$$

In the multiplier model, the focus is on the weighting variables, μ and ν, which are the "prices" of the inputs and outputs. The weightings (μ and ν) are the dual variables relating to the constraint in the envelopment model (Equation 2 - 5), corresponding to inputs and outputs. In this model, the μ can be interpreted as the imputed marginal value or shadow price of outputs. The ν can be interpreted as the imputed marginal value or shadow price of inputs. These prices are determined to maximise a DMU's efficiency score. The multiplier form provides prices directly from the weights from multipliers.

Mathematical programming duality ensures that each model can provide efficient quantity combinations or prices for the other model. By solving either the envelopment model or the multiplier model, researchers can determine physical quantities or prices. Based on duality

[8] In an output-oriented model, λs are the expansion in outputs for a DMU to be technically efficient.

theory, the optimal solutions for the efficiency scores are the same for both models (Shephard, 1970).

In the primal space (quantity space), the technical efficiency measured by DEA follows the definition by Farrell (1957) (section 2.2.1) and corresponds to the input and output distance functions (Färe & Primont, 1995; Shephard, 1970; Thanassoulis, 1997). In the quantity space, the DMUs within a production possibility set under a certain technology are compared with relatively efficient DMUs forming the production frontier. The distance function can be used to calculate the relative efficiency scores for each DMU, as discussed in section 2.2.1. There are two types of radial distance functions[9]: an input distance function, which represents a minimal proportional contraction of the inputs for given outputs, and an output distance function, which represents a maximal proportional expansion of outputs for given inputs (Coelli et al., 2005). In economic duality, the input and output distance functions are related to the cost and revenue functions, respectively. The relationship between input or output distance functions and cost or revenue functions means they both contain identical information about the production technology and each of them can be obtained from the other. In other words, there is a one-to-one relationship between the input distance function and the cost function, and between the output distance function and the revenue function (Färe & Primont, 1995; McFadden, 1978).

The duality relationship, as explained above, can be used to estimate the input and output prices. The prices are known as imputed marginal values, shadow prices or virtual prices for inputs and outputs of DMUs (Thanassoulis, 1997). Of particular interest is the situation where the market prices might be unknown for a subset of inputs and outputs but known for the other inputs and outputs. The relativities among the shadow prices can allow these market prices to be imputed to those inputs and outputs without market prices, e.g. pollution, quality, capital.

From duality theory, prices can be derived from the quantities which reflect the opportunity costs of resources in terms of the usual linear programming models. If price information is known about inputs and outputs in addition to quantities, allocative efficiency can be calculated using revenue, cost, or profit models where the decision variables are the quantities for the DMU under evaluation to optimise to achieve allocative efficiency.

[9] A radial efficiency is defined as when all inputs can be simultaneously reduced without altering the mix or proportions in which they are utilised. Or when all outputs can be simultaneously increased without altering the mix or proportions they are produced (Cooper et al., 2007a).

In production models, physical data is used to generate TE scores. However, in Financial DEA, accounting variables are essentially the product of prices and quantities (influenced by accounting choices and regulations). When the only information given is the accounting variables, but not separate prices or quantities, then in order to use accounting data to calculate TE, certain conditions need to be met (Färe et al., 2017). For example, if the market prices of inputs and outputs are the same across all DMUs, accounting data can be used to estimate TE (Cross & Färe, 2009; Färe & Grosskopf, 1985; Färe et al., 1990). By contrast, when prices are unknown and different across DMUs, using accounting data to estimate TE can be problematic (Portela, 2014). In other words, when the uniformity of prices starts to vary, using Financial DEA to estimate TE becomes problematic. For example, when prices vary around a narrow band (i.e. in a competitive market), prices across firms are similar. In this case, Financial DEA can use accounting data to estimate and interpret TE with a moderate level of measurement error. However, when prices vary around a broad band (i.e. in a less competitive market), all firms face different prices. In this case, Financial DEA results using accounting data may distort the interpretation of TE, and the efficiency measures will be a mix of technical and allocative efficiency with no clear-cut distinction. In Financial DEA, where accounting variables are confounded by quantities and different prices (and impacted by various accounting choices and regulations), neither qualities nor prices can be separated. In this case, the multiplier weights may not directly relate to prices.

2.2.2.3. Banker, Charnes and Cooper (1984): BCC model

Banker et al. (1984), based on CCR (1978), further isolated the scale efficiency and therefore distinguished DEA model choices to constant returns to scale (CRS or CCR) (Charnes et al., 1978), or variable returns to scale (VRS or BCC) (Banker et al., 1984).

The difference between CRS and VRS DEA models is whether the economies of scale effect is incorporated in the DEA modelling. CRS assumes all DMUs operate at an optimal scale. Mathematically, CRS assumes a proportional change in inputs and outputs. For instance, an increase (decrease) in inputs has the same proportional size increase (decrease) in outputs. When the inputs and outputs are not scaled perfectly, VRS is a more suitable choice. VRS adds increasing returns to scale (IRS) and decreasing returns to scale (DRS) to CRS. IRS represents a more-than-proportional change in outputs to the change of inputs. DRS represents a less-than-proportional change in outputs to the change of inputs. Therefore, scale efficiency (SE), which is the efficiency due to the economies of scale effect, can be obtained

by decomposing the CRS TE to the combination of VRS TE ("pure TE") and SE (Coelli et al., 2005). Also, DEA models are based on a range of axioms taken from the economic production perspective, which act as the assumptions of DEA, which are explained next.

2.2.2.4. DEA axioms

The construction of the DEA model is based on the following axioms (Banker et al., 1984; Banker et al., 1993; Shephard, 1970):

(1) Inclusion of observations
(2) Free disposability or monotonicity
(3) Convexity of the production function
(4) Constant returns to scale
(5) Minimum extrapolation

First, the inclusion of the observation axiom assumes an efficient frontier is an empirical frontier and includes all observations. An efficient frontier is also known as a production possibility frontier. Unlike a theoretical frontier based on pure theory, the DEA efficient frontier is based on observed DMUs and is technologically feasible. All inputs and outputs of the DMUs belong to a production possibility set:

$$T = \{(X, Y) | Y \geq 0 \; can \; be \; produced \; from \; X \geq 0\} \qquad (2-6)$$

Within the production possibility set, the input possibility set is:

$$L(Y) = \{X | (X, Y) \in T\} \qquad (2-7)$$

Similarly, the output possibility set is:

$$P(X) = \{X | (X, Y) \in T\} \qquad (2-8)$$

Intuitively, the inclusion of an observation axiom states that if a DMU has produced a certain volume of outputs with a certain volume of inputs, this DMU would be capable of repeating this production action with the same quantities in the future while holding everything the same.

Second, the free disposability or monotonicity axiom is stated as:

$$(a) \qquad If\ (X,Y) \in T, and\ \overline{X} \geq X, then\ (\overline{X},Y) \in T$$

$$(b) \qquad If\ (X,Y) \in T\ and\ \overline{Y} \leq Y, then\ \left(X,\overline{Y}\right) \in T \qquad (2-9)$$

This axiom assumes that a DMU can keep producing the same amount of output with the same or more input. Similarly, a DMU can keep producing the same or fewer outputs with the same amount of input. This axiom is also called the "inefficiency postulate" in Banker et al. (1984) to indicate the possibility of inefficient production with more inputs, fewer outputs, or both.

Third, the convexity of the production function states:

$$If\left(X_j,Y_j\right) \in T, and\ \lambda_j \geq 0\ are\ nonnegative\ scalars\ such\ that\ \sum_{j=1}^{n} \lambda_j = 1,$$

$$then\ \left(\sum_{j=1}^{n} \lambda_j X_j\ , \sum_{j=1}^{n} \lambda_j Y_j\right) \in T \qquad (2-10)$$

This axiom assumes that new DMUs can be a convex combination of two observed DMUs. By combining two observed DMUs convexly, a new DMU is also feasible. Graphically, the convex axiom further changes an efficient frontier from a jagged hull to a piece-wise linear frontier.

Fourth, the constant returns to scale axiom states:

$$If\ (X,Y) \in T, then\ (\lambda X, \lambda Y) \in T, \lambda \geq 0) \qquad (2-11)$$

This axiom changes the efficient frontier to the shape of a straight line. As a result, efficiency scores for the same DMU have higher VRS efficiency scores (or the same, if on the CRS part of a VRS frontier) than CRS efficiency scores.

Fifth, the minimum extrapolation states that *T* is the intersection set of all $\hat{T}$ satisfying axioms (1) to (4) and subject to the condition that each of the observed vectors $(X,Y) \in \hat{T}$. This axiom estimates the smallest function that envelopes all observed DMUs. Since *T* is based on a convex and ray extension, *T* is a polyhedral set (Banker et al., 1984).

2.2.3. Alternative Views of DEA

Cook et al. (2014, p. 2) stated that "although DEA has a strong link to production theory in economics, the tool is also used for benchmarking in operations management". When DEA represents an underlying production process, inputs are the resources used to generate outputs (Cook et al., 2014, p. 2). For example, in Financial DEA literature, researchers used the cost of labour to proxy the labour element; the cost of material to proxy the material element; and the cost of capital to proxy the capital element according to the underlying manufacture processes in the dairy and food industries (Aparicio & Kapelko, 2018; Kapelko et al., 2016). This view treats accounting measures as proxies of production elements.

Alternatively, DEA can solve benchmarking problems where the inputs are the "less-the better", and the outputs are the "more-the-better" performance measures. This benchmarking case is relevant to the situations where DEA is used as a Multiple Criteria Decision Making (MCDM) tool (Cook et al., 2014, p. 2). MCDM has been a discipline in its own right since the 1970s (Stewart, 1992). In the 1990s, researchers started to comment that DEA can be used as an MCDM tool if the inputs and outputs in DEA are viewed as attributes or criteria to evaluate DMUs (Belton & Vickers, 1993; Doyle & Green, 1993; Stewart, 1996). In more recent years, it has become common practice that researchers use DEA as an MCDM tool to examine practical issues (da Silva et al., 2021; Martín-Gamboa et al., 2017).

2.2.4. Methodological Issues in DEA

There are various methodological issues to be considered when applying DEA. This section reviews four key issues, being homogeneity, discriminatory power, non-negative variables, and selection of variables.

First, DEA assumes DMUs under assessment are homogenous in several ways. Business activities are homogeneous in that they utilise similar resources to produce comparable goods and services. Also, technologies are homogenous, although this can be varied depending on the method used. Operating environments should also be similar since the external environment could impact the overall performance of DMUs (Dyson et al., 2001; Golany & Roll, 1989).

Second, the discriminatory power is the ability of DEA to distinguish between the efficient and inefficient DMUs. The relationship between sample size (number of DMUs) and model

size (number of inputs and outputs) can influence the discriminatory power of DEA models. Discriminatory power is low when efficiency scores are spread in a narrow range around full efficiency.[10] Low discriminatory power is usually due to the inclusion of an excessive number of variables in a DEA model (Cook et al., 2014). Researchers need to ensure sufficient discriminatory power by finding a balance between a comprehensive DEA model and reasonably sized sample sets.

Researchers have various views of the "rule of thumb" to decide the minimum number of DMUs to achieve DEA results with reasonable discriminatory power. For example, Golany and Roll (1989) suggested the number of DMUs should be no less than $2 \times (m + s)$, where m is the number of inputs and s is the number of outputs. The minimum number of DMUs under this view is six ($2 \times (2 + 1)$). Banker et al. (1989) and Cooper et al. (2007b) suggested that the number of DMUs be greater than the maximum of $3 \times (m + s)$ and $m \times s$. The minimum number of DMUs under this view is nine ($3 \times (2 + 1) > 2 \times 1$). Dyson et al. (2001) suggested that the number of DMUs should be no less than $2 \times m \times s$. The minimum number required under this view is four ($2 \times 2 \times 1$).

Third, DEA requires non-negative variables to generate feasible results. As suggested by researchers, there are two common ways to treat negative variables. The first is that if the DMUs with negative variables cover a small portion of the data set, remove those DMUs from the group (Bowlin, 1995). The second is to translate the negative values by adding a constant to the variable for each DMU equal to at least the largest negative value in the sample (Ali & Seiford, 1990; Bardhan et al., 1996; Bowlin, 1998; Bowlin, 1999).

The fourth key methodological issue is that the selection of variables for DEA models determines factors counted for efficiency. DEA can be used under alternative views of productive efficiency or benchmarking (Cook et al., 2014). Under the view of productive efficiency, based on the underlying production function, variables are chosen to proxy the physical productive elements (Boussofiane et al., 1991; Golany & Roll, 1989). Under the view of benchmarking, variables are chosen to build performance metrics according to the performance goal (Belton & Vickers, 1993; Doyle & Green, 1993; Stewart, 1996).

[10] This could also happen in groups of DMUs that have high structural efficiency. That is homogeneous sets with very similar levels of productivity.

2.2.5. Summary

To summarise, this section examines the fundamentals and key features of DEA. The origin of DEA and its linear programming formula suggest that, from an economic viewpoint, it is necessary to distinguish price variables from physical variables in DEA research. The physical form and monetary form of production functions are only transformable under certain conditions. In more recent years, researchers have used DEA as an MCDM tool, through which DEA models evaluate alternative DMUs according to a performance goal. This section also reviews four key methodological issues to be considered when applying DEA. The next section will introduce the nature of accounting information.

2.3. Accounting Information

Accounting information is the language of business as it enables organisational participants to communicate and work as a whole. However, researchers have inconsistent views of the usefulness of accounting information. On the one hand, accounting information has been viewed as useful for decision making since it enables participants' cooperation, resource allocation, and organisation function (Jordan & Messner, 2012). On the other hand, accounting information has been viewed as having limited usefulness because of its heterogeneous nature, given that it is a convenient way of summing up many different components (Ball & Brown, 2014; Williams, 2014). For example, accounting numbers aggregate quantities and prices. Also, accounting numbers aggregate across activities like production and administrative activities.

Ball and Brown (1968) espoused the usefulness of accounting information as a relevant signal of performance by decision-makers since their random walk model empirically supported that share prices reacted to accounting information. Specifically, good (bad) accounting information[11] impacted share prices positively (negatively). Similarly, professional accounting bodies explicitly define the qualitative features of accounting information to assist its usefulness.

According to the International Accounting Standards Board's ([IASB]'s) conceptual framework for financial reporting, "if financial information is useful, it must be relevant and

[11] Good accounting news means the accounting information conveying good news. Similarly, bad accounting news means the accounting information conveying bad news.

faithfully represent what it purports to represent. The usefulness of financial information is enhanced if it is comparable, verifiable, timely and understandable" (IASB, 2018, para 2.4). The fundamental qualitative characteristics of financial information are relevance and faithful representation. Relevance is defined as "relevant financial information (which) is capable of making a difference in the decisions made by users" (IASB, 2018, para 2.6 - 2.10). Faithful representation means "representation of the substance of an economic phenomenon instead of representation of its legal form only" (IASB, 2018, para 2.12).

However, in practice, the qualitative characteristics of the conceptual framework are impacted by various contextual factors. Therefore, the qualitative characteristics may not be perfectly represented in practice. For example, Barth et al. (2012) stated that "comparability of accounting information is a function not only of accounting standards, but also of interpretation, auditing, and the regulatory, litigation, and enforcement environment" (p. 70). McCallig et al. (2019) found that auditing opinions and stakeholders who need credible information can enhance the faithful representation of accounting information. Barth et al. (2021) found that the relevance of accounting information for assessing firm value became more nuanced with the development of information technology and growth in intangible assets. These contextual factors may impact the qualitative characteristics of accounting information to various degrees. The accounting information created under different contexts may vary in qualitative characteristics. And this heterogeneity may introduce measurement errors to Financial DEA through accounting information.

Also, researchers have questioned the usefulness of accounting information as it does not naturally represent firm performance and contains heterogeneous and incomplete information (Chua, 1995). Even Ball and Brown (1968) acknowledged their counterargument that the heterogeneous nature of accounting information could impair its usefulness (Chambers, 1964). In particular, accounting information aggregates various business areas (e.g. taxation, leasing, revenue) and value estimations (e.g. historical cost and inflations).

Even though accounting information is quantified, it also has qualitative features open to different interpretations. Quantitative accounting numbers are not equivalent to physical reality as measured in the natural sciences. In the natural sciences, measurements are objective and involve the physical operation of comparison of the quantity being measured with some standard (Stamp, 1993). For example, a concept such as time has quantity measures by nature. Physicists can use these measures mathematically to model the physical

world and make predictions. By contrast, in accounting, quantity can be ambiguous since the choice of metric is arbitrary and subject to various measurement applications, for example, the arbitrarily chosen property, such as cost or fair value. Accounting numbers are not quantities compared to physical reality but are only numerical representations of interpretative concepts and often involve disagreement. Accounting numbers can be problematic since it can be ambiguous as to what is represented (Andon et al., 2015; Williams, 2014).

As discussed in section 2.2.2.2, the duality theorems between the physical space and price space are built on a range of assumptions. When using accounting information in DEA, these assumptions are not necessarily met. For example, accounting information aggregates various prices occurring over different time periods with various physical quantities. As a result, accounting data may not estimate the physical TE as duality theorems state.

2.4. Firm Performance

2.4.1. Overview of Firm Performance

Firm performance is one of the most important concepts in organisational research. Despite its importance, the literature has developed many perspectives on the definition and measurement of firm performance (Miller et al., 2013). Theoretically, researchers have treated firm performance as (a) a specific construct, (b) an aggregated construct, and (c) a general construct (Hamann & Schiemann, 2021; Miller et al., 2013; Richard et al., 2009). This section will introduce these treatments of firm performance, with examples from the literature on Financial DEA.

2.4.2. Firm Performance – Specific Construct

Under the view of a specific construct, firm performance is viewed as loosely related and separate constructs that are conceptually distinct. Researchers do not conceptualise firm performance in a general way. Instead, researchers use specific constructs, such as profitability or marketability and provide theories specifically linked to the particular constructs. More than one dimension of firm performance is incorporated in research, and separate conceptualisation and theories are expected. Researchers accept the trade-off between simplicity and accuracy under the view of specific constructs. Researchers develop

distinct arguments for specific performance aspects to provide meaningful grounding for specific performance variables (Hamann & Schiemann, 2021; Miller et al., 2013; Richard et al., 2009).

Researchers have provided conceptual frameworks to serve this purpose to measure firm performance under a specific construct view. For instance, Porter (1985) proposed a value chain to describe dimensions within a firm based on activities (Figure 2 - 2). The value chain is a set of activities carried out by a firm to deliver valuable products to the market. These activities are defined as building blocks by which a firm creates a valuable product for customers. The margin is defined as the difference between the total value and the collective cost of performing the value activities.

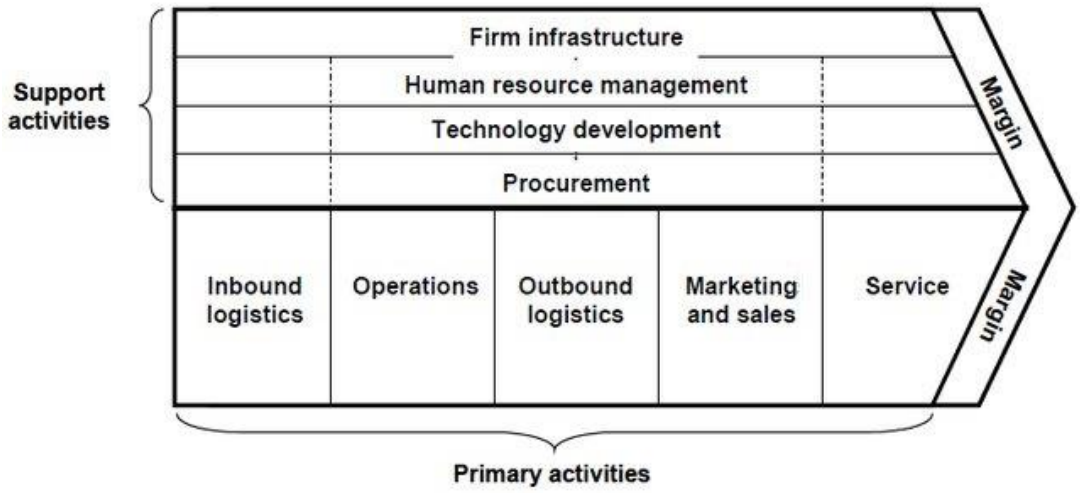


Figure 2 - 2 Porter's Value Chain from Porter (1985)

Similarly, Venkatraman and Ramanujam (1986) proposed that firm performance could be defined as three alternative constructs with distinct domains. As illustrated in Figure 2 - 3, the three constructs are represented by concentric circles overlapping each other, indicating the different scope of each domain: (a) organisational effectiveness (the largest circle), which relates to various organisational goals, (b) business performance (the medium circle), which is the combination of financial and non-financial (operational) performance, and (c) financial performance (the narrowest circle), which reflects the economic goals.

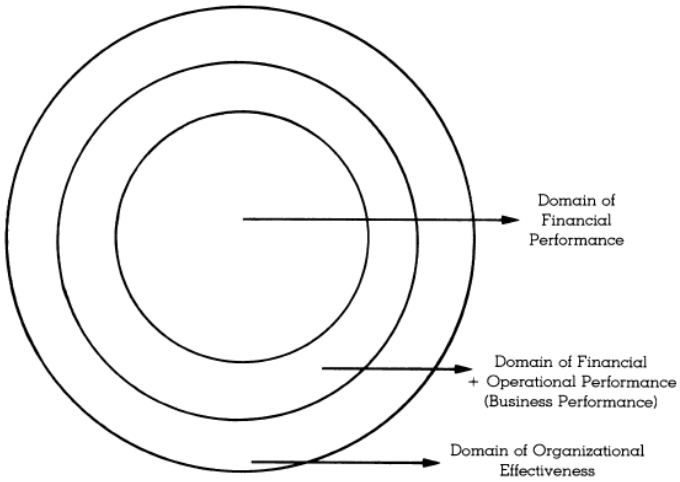


Figure 2 - 3 The Domain of Firm Performance from Venkatraman and Ramanujam (1986)

In more recent years, researchers have proposed various dimensions of firm performance. For example, Combs et al. (2005) suggested that firm performance covers four dimensions: accounting returns, growth, stock market and operational performance. Richard et al. (2009) suggested three dimensions of firm performance: financial performance, shareholder return, and product market performance. Hamann and Schiemann (2021) suggested four dimensions of firm performance: profitability, liquidity, growth, and stock market performance. Hamann and Schiemann (2021) also used confirmatory tetrad analysis (CTA) to empirically confirm that these four dimensions are separated and not interchangeable.

In Financial DEA, several empirical studies exemplify the three separate dimensions of firm performance. For example, to measure financial performance, Edirisinghe and Zhang (2008) used "an approach to combine financial statement data using Data Envelopment Analysis to determine a relative financial strength" (p. 842). To measure the operational performance, Min and Joo (2006) suggested: "a Data Envelopment Analysis (DEA) that is proven to be useful for measuring the operational efficiency of various profit or non-profit organisations" (p. 259). To measure effectiveness, Tsolas (2011) used "Data Envelopment Analysis (DEA) … to evaluate performance in terms of profitability and effectiveness" (p. 795).

2.4.3. Firm Performance – Aggregated Construct

Firm performance has been treated as an aggregated construct with multiple components. Specifically, firm performance is viewed as a holistic construct with disparate dimensions, and these divergent dimensions combine to define the overall firm performance (Hamann & Schiemann, 2021; Miller et al., 2013). Compared to the specific construct view, the aggregated construct view combines individual constructs with an algebraic function (Hamann & Schiemann, 2021; Law et al., 1998).

For example, the three-E performance framework is used to aggregate firm performance, especially in public organisations. According to the three-E performance framework, three disparate dimensions are components to aggregate the firm performance construct, and these are (a) efficiency, (b) effectiveness, and (c) economy (Harrison, Rouse, & De Villiers, 2012; Ramanathan, 1985). As demonstrated in Figure 2 - 4, efficiency is defined as the ratio of outputs to inputs to maximise productivity by increasing outputs or decreasing inputs. Effectiveness is defined as the ratio of outcomes to outputs to achieve organisational goals. Economy is defined as the relationship between inputs and outcomes, which is concerned with the specific control issues related to the public sector.

Figure 2 - 4 Three-E Performance Framework from Rouse (2006)

Several Financial DEA studies empirically exemplify the three-E performance framework. For instance, Athanassopoulos (1995) proposed a Financial DEA model for local governments that aggregated "the objectives of resource management, namely effectiveness, equity and efficiency" (p. 543). Similarly, Doumpos and Cohen (2014) presented a Financial DEA model to measure the "efficiency, effectiveness, and economy considerations in the public sector" (p. 74).

Similarly, the balanced scorecard views firm performance as a construct with four integral perspectives: (a) the financial perspective is mainly concerned with the financial results for the shareholders; (b) the consumer perspective is related to creating value for customers and retaining them; (c) the internal business process focuses on the practices within businesses, such as firm operational efficiency; and (d) the learning and growth perspective is to improve performance (Kaplan & Norton, 1992).

DEA studies empirically exemplify that a DEA model can aggregate these four performance aspects in the balanced scorecard into a single score that holistically measures the firm's performance (Amado et al., 2012; Chen, T. et al., 2008). For example, Rouse et al. (2002), as illustrated in Figure 2 - 5, used the "DEA model to reflect … strategy provides a balanced view of organisational performance" (p. 229). The authors selected performance data within each perspective of the balanced scorecard to measure firm performance.

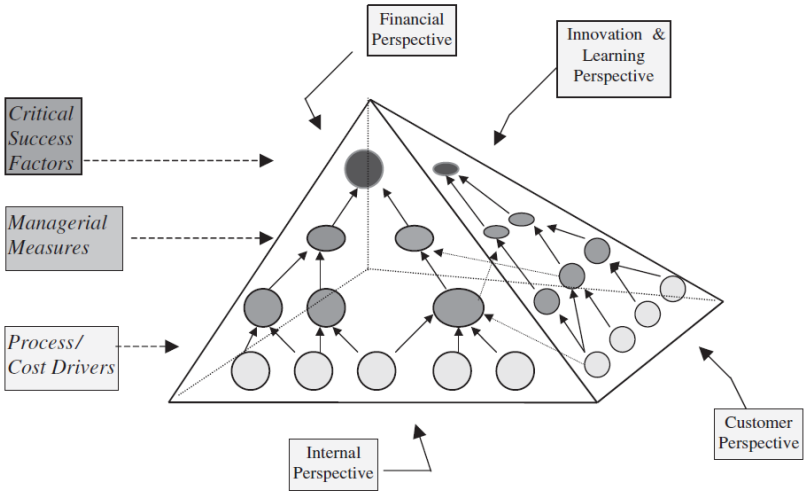


Figure 2 - 5 Integrated BSC-DEA Performance Measurement Model from Rouse et al. (2002)

Another common conceptualisation of aggregated firm performance is the decomposition of the DuPont ratio. The DuPont company developed this ratio in the 1920s (Soliman, 2008). By decomposing the ratio of return on equity (ROE), the DuPont ratio aggregates three dimensions into the construct of firm performance, being (a) the profitability efficiency, (b) the asset utilisation efficiency, and (c) the financial leverage, or equity multiplier.

Financial DEA studies exemplify the aggregated firm performance construct through DuPont ratio decomposition. For instance, Feroz et al. (2003) stated: "DEA can be applied to revenue-producing organisations… one such approach is to disaggregate Return on Equity (ROE) using the DuPont model" (p. 49). The authors empirically built Financial DEA models upon the DuPont ratio that the Financial DEA model variables were selected from, based on the three dimensions in the DuPont ratio.

2.4.4. Firm Performance – General Construct

Firm performance as a general construct is assumed to be a general phenomenon with an implicit meaning. This definition of firm performance is relatively abstract in research. Researchers neither describe different performance dimensions nor specify the relationship between one dimension of performance with other variables. Instead, the definition of firm performance is parsimonious, and the focus is to build a logical relationship between firm performance with other variables. Even if a researcher defines multiple variables to measure firm performance, these variables would be used to minimise measurement errors of each other instead of expanding on the number of features of firm performance (Richard et al., 2009). For example, Spanos et al. (2004) used price-cost margin to measure the construct of profitability representing firm performance. Hillman and Keim (2001) used the market value-added to measure shareholder value creation, representing firm performance. In recent years, Tobin's Q has been used commonly to measure the construct of accounting performance and to represent firm performance (Bennouri et al., 2018; Cappa et al., 2021; Girod & Whittington, 2017).

In Financial DEA, Chang et al. (2013) exemplifies the general construct view. The authors focused on measuring the relationship between firm diversification and firm performance. In theory building, firm performance was defined generally, without specifying dimensions. In their empirical modelling, firm performance was measured by Financial DEA models with alternative variables. In the primary model, they incorporated as inputs the cost of goods sold (*COGS*), selling and distribution expenses (*XSGA*), capital expenditure (*CAPEX*), and sales revenue as an output. In the sensitivity test, the authors replaced *CAPEX* with depreciation expenses, net assets, and the opportunity cost of capital to improve robustness. The authors used alternative measures of firm performance to minimise measurement errors. Since the authors defined firm performance parsimoniously without specifying its dimensions, the key focus was the relationship between firm performance and firm diversification.

In summary, the literature on firm performance suggests various views of the construct of firm performance that can be used to build Financial DEA models. Under different views, the construct Financial DEA models are built differently. The next section will review the literature of measurement models, which can explain how the construct of firm performance can be modelled in the Financial DEA application.

2.5. Measurement Models

Measurement models are defined as the models that quantify the associations between observations obtained during research and the underlying theoretical factors (American Psychological Association, 2020). This section introduces the fundamental definitions used in the general research process and measurement models to provide a basis for examining the measurement models used in Financial DEA.

2.5.1. Empirical Research Process

The general empirical research process describes steps to carry out empirical research, as illustrated in Figure 2 - 6. Applying this process, problems are refined to generate clear and testable research questions.

In the general research process, analyses exist at both the conceptual and operational levels (Bisbe et al., 2007). At the conceptual level, theory (A) identifies the constructs of interest and clarifies their meaning. Constructs (B and C) are theoretical creations defined in conceptual terms but are not directly observable (Edwards & Bagozzi, 2000). The process of specifying constructs based on the theory is known as conceptualisation, where imprecise theories are refined to the precise and agreed-upon meaning of constructs. The theory also specifies conceptual relationships (Link 1) between constructs.

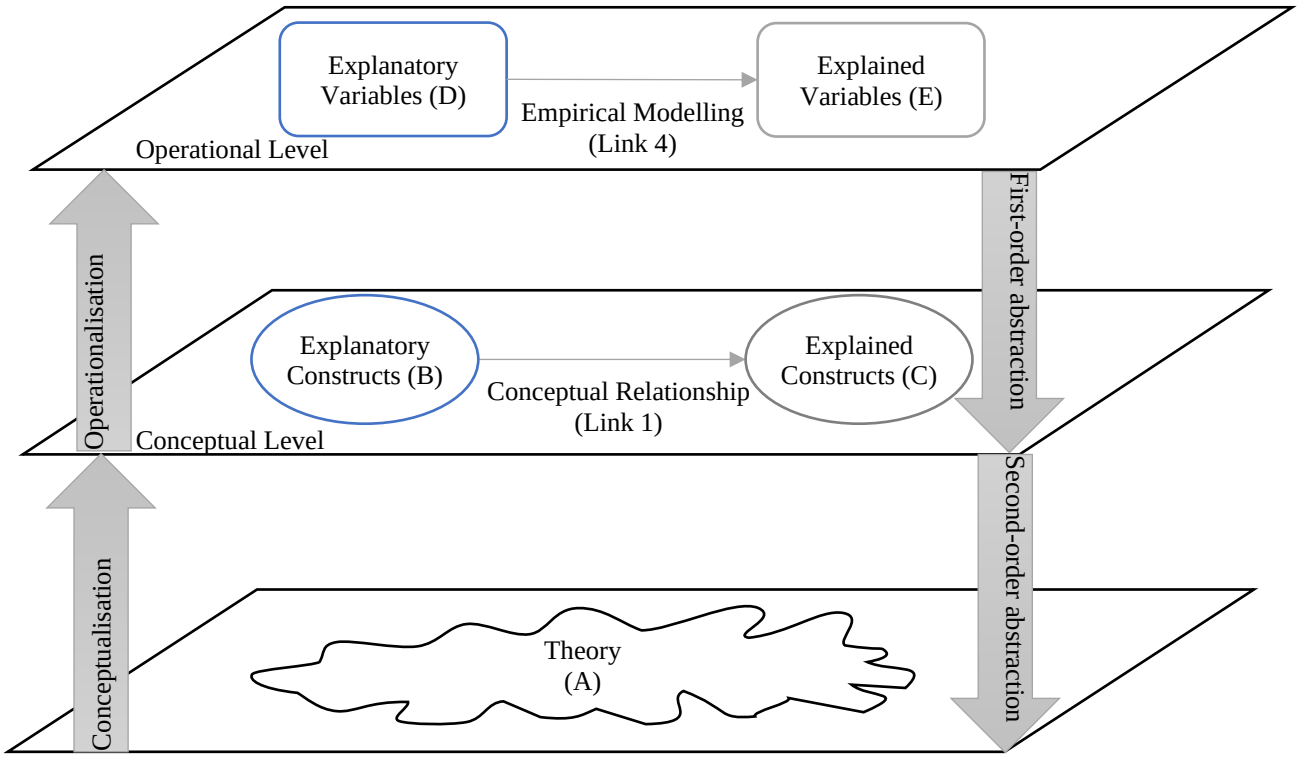


Figure 2 - 6 General Empirical Framework
Adapted from Ghauri and Grønhaug, 2010, Bisbe et al., 2007, and Mackenzie et al., 2005

For example, the Financial DEA study by Demerjian et al. (2013), which was based on shareholder theory[12] (A), specified the relationship between the managerial ability construct (B) and the earnings quality construct (C). The authors expected that more capable managers would better know their firms, leading to better judgements, estimations, and higher quality earnings (Link 1).

Since constructs are not observable, they anchor to the observable reality through indicators. Indicators are defined as signs of the presence or absence of a construct. This is done by referring to its observable manifestations (when indicators reflect underlying latent constructs in a reflective model) or constitutive facets (when indicators are aspects of forming a composite construct in a formative model).[13] Operational definitions are applied to measure indicators, which are defined as empirical referents that specify operations to be carried out.

[12] That is, high-quality earnings accurately reflect companies' operating performance, and the managers take a shareholder perspective when determining the desired attributes of earnings.
[13] Section 2.5.2 details the differences between manifestations and facets, and formative and reflective measurement models.

The results are known as operational variables. The observed score of an operational variable is defined as a measure (Bisbe et al., 2007; Edwards & Bagozzi, 2000). Research then moves from the conceptual level to the operational level, which is known as operationalisation. Operationalisation is defined as the process that translates constructs (B and C) into operational variables (D and E) (Babbie, 2016; Bisbe et al., 2007). The relationship between constructs is subsequently tested by empirical models (Link 4), from which the theory (A) is indirectly tested.

For example, in Financial DEA studies, Demerjian et al. (2013) operationalised the construct of managerial ability (B in Figure 2 - 6) using the residuals from efficiency scores. This financial efficiency was generated by a Financial DEA model (D in Figure 2 - 6). In this model, inputs were cost of goods sold, selling and administrative expenses, net property plant and equipment, net operating leases, net research and development, purchased goodwill, and other intangible assets. And the output was sales. Demerjian et al. (2013) operationalised the construct of earnings quality (C in Figure 2 - 6) using earnings restatement, earnings persistence, the provision for bad debts, and erroneous accruals (E in Figure 2 - 6).

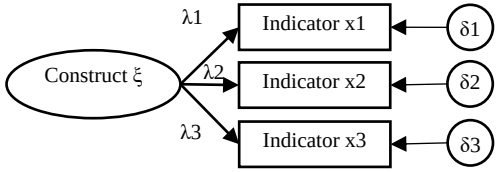
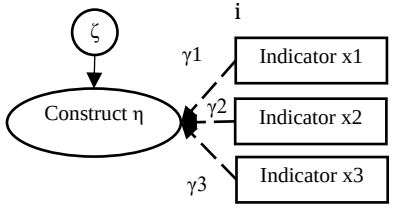
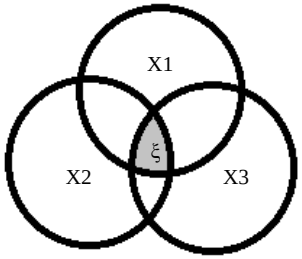
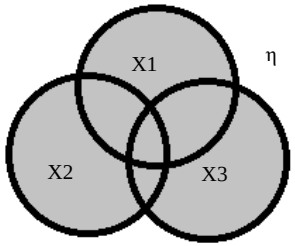
This thesis focuses on the conceptualisation and operationalisation processes. The explanatory and explained construct-and-variable sets are two parallel processes that follow the same methodologic pathway. This study uses the explained construct-and-variable set as the example used in the rest of this chapter (refer to elements coloured in blue Figure 2 - 6).

2.5.2. Reflective and Formative Models

This section discusses how different measurement models impact the relationship between variables and constructs based on the general empirical research process. Measurement models bridge the abstract theoretical constructs and measurable empirical variables. Without measurement models, the mapping of theoretical concepts onto empirical models is ambiguous, and theories cannot be meaningfully tested (Blalock, 1971; Edwards & Bagozzi, 2000).

This section introduces two types of measurement models: (a) the reflective model, where indicators are manifestations of an underlying latent construct, and (b) the formative model, where indicators are constitutive facets of a composite construct (Bisbe et al., 2007; Hamann & Schiemann, 2021; Jarvis et al., 2003). The key features of the two models are summarised in Table 2 - 1.

Table 2 - 1 Measurement Models[14]

	Reflective Model	**Formative Model**
The direction of Causality/ Change	from constructs to indicators	from indicators to constructs
Relationship Between Indicators	correlated, covariant and interchangeable	not necessarily correlated, covariant or interchangeable
Result of Dropping Any Indicators	the domain of construct does not change	the domain of construct changes
Measurement Error	at measurement level (δ)	at the construct level (ζ)
Nomological Net of Indicators	same antecedents and consequences in the same nomological net	not necessarily same antecedents and consequences or same nomological net
Equation	$x_i = \lambda_i \xi + \delta_i$	$\eta = \sum_i \gamma_i x_i + \zeta$
Diagram Side view		
Diagram Top-down view		

[14] This table is adapted from Jarvis et al. (2013), Edwards and Bagozzi (2000), Law and Wong (1999).

2.5.2.1. ***Reflective model***

In a reflective relationship, indicators reflect the underlying latent construct. The direction of causality is from the constructs to the indicator. A change of the underlying construct leads to a change in the indicators. Since all indicators share the same cause (i.e. the underlying construct), indicators are supposed to co-vary, and the covariation attributes to the common cause – the construct. Any two of the indicators correlate, and they are interchangeable. All indicators should share the same nomological net[15] (i.e. same antecedents and consequences). Even dropping an indicator would not change the domain of interest, as the remaining indicators adequately represent the construct (Bisbe et al., 2007; Jarvis et al., 2003).

Graphically, a reflective construct is the common area of different indicators (Table 2 - 1, the top-down view diagram). In the side-view diagram, the underlying construct (ξ) is measured by indicators (x_i) with the influence of random measurement errors (δ_i) (Edwards & Bagozzi, 2000). The random measurement errors are due to the invalidity and unreliability of the individual measures, which could be caused by imperfect measures or random factors (Mackenzie, Podsakoff, & Jarvis, 2005). For each variable (x_i), the corresponding λ_i represents the loading factor of the effect of the underlying construct (ξ) on the indicator (x_i) (i ranges from 1 to 3 for the example shown) (Edwards & Bagozzi, 2000). The unique error explains the variance in indicators (x_i), and covariation among indicators is attributed to the common construct (ξ) (Bisbe et al., 2007; Edwards & Bagozzi, 2000; Jarvis et al., 2003; Law et al., 1998; Law & Wong, 1999).

If variables operationalise an incomplete set of indicators in a reflective model, the model's reliability would be reduced, but not necessarily its validity. Reliability is defined as the stable and consistent accuracy of measurements, and validity is the degree that the measurements capture what they are supposed to measure (Ghauri & Grønhaug, 2010). In reflective models, indicators reflect the common underlying latent construct, and variables are the operationalised indicators. Therefore, in a reflective model, when not all the indicators are operationalised to variables, the variables can still reflect the construct, but they are not stable (Bisbe et al., 2007).

[15] Nomological net is a representation of the concepts or constructs of interest in a study, their observable variables, and the interrelationships among them (Preckel & Brunner, 2017).

2.5.2.2. ***Formative model***

In a formative relationship, indicators form a composite construct by describing their inherent constitutive facets. Indicators are different facets and jointly work as a group to determine the conceptual meaning, which is the relevant domain (Bisbe et al., 2007; Jarvis et al., 2003). The direction of causality is from indicators to the construct. Therefore, a change in an indicator leads to a change in the relevant domain of the construct. Every indicator plays a unique role in forming the definition of the construct. Any two indicators are not necessarily correlated or interchangeable. Indicators do not necessarily have the same antecedents or consequences (i.e. the same nomological net). Therefore, an essential feature of formative models is that dropping a constitutive facet would change the domain of the construct. Researchers need a census of indicators instead of their sample (Bisbe et al., 2007; Jarvis et al., 2003).

In a formative model, indicators jointly form a composite construct by describing its inherent attributes. The causality direction is from the indicators to the construct. Therefore, any two indicators are not necessarily interchangeable or correlated (Bisbe et al., 2007; Edwards & Bagozzi, 2000; Jarvis et al., 2003). For example, Reilly (1982) measured the construct of family social status with Warner's (1949) Index of Status Characteristics, which was composed of the rating of occupation, source of income, dwelling type, and neighbourhood quality. The four aspects are not interchangeable, nor do they necessarily co-vary with each other. Therefore, the indicators should be viewed as formative indicators of social status (Jarvis et al., 2003).

Graphically, a formative construct is the total area of different indicators (Table 2 - 1, the top-down view graph). In the side-view graph, the underlying construct (η) is formed by indicators (x_i) and disturbances (ζ). A disturbance is defined as the part of the construct (η) that cannot be explained by indicators (x_i) (Edwards & Bagozzi, 2000). The disturbance captures the construct's invalidity, which may be due to the imperfect validity of the individual facets or their invalidity as a group due to failing to include all the construct facets (Mackenzie et al., 2005). The indicators (x_i) are defined as error-free causes of the construct (η) since the direction of causality is from the indicators to the construct (Maccallum & Browne, 1993) (i ranges from 1 to 3 for the example shown). The disturbances exist on constructs (Bisbe et al., 2007; Edwards & Bagozzi, 2000; Jarvis et al., 2003; Law et al., 1998; Law & Wong, 1999).

Construct validity is defined as the extent to which a variable correctly represents the relevant domain that was meant to be measured (Bisbe et al., 2007; Ghauri & Grønhaug, 2010). Since the domain of the construct is not adequately covered by omitted indicators, the meaning of the composite construct changes. If variables operationalised an incomplete set of indicators in a formative model, the construct validity would be undermined. The deterioration of construct validity could also be caused by including imperfect variables (Bisbe et al., 2007; Ghauri & Grønhaug, 2010).

2.5.2.3. Relationship between constructs and phenomenon

As illustrated in Figure 2 - 7, the discussion so far only focuses on the relationship between variables and constructs, which is named the first-order model in the first-order abstraction. However, measurement models can expand to the more theoretical abstract – the second-order model, which is in the second-order abstraction and covers the relationship between constructs and phenomenon (Bisbe et al., 2007; Jarvis et al., 2003).[16] For simplicity, this study names first-order models as the relationships between variables and constructs. Similarly, second-order models as the relationship between constructs and phenomenon.

In Financial DEA, the first-order constructs are dimensional constructs of firm performance phenomenon. These dimensional constructs are measured by Financial DEA models, where the input and output accounting variables are operationalised indicators. Further, the second-order construct is firm performance phenomenon (section 2.4). For example, Demerjian et al. (2012) and Demerjian et al. (2013) used seven inputs and one output in a Financial DEA model to measure the dimensional construct of financial efficiency at the first-order abstraction level. Financial efficiency was used at the second-order abstraction level to represent firm performance phenomenon since the researchers treated firm performance as a general phenomenon without specifying any dimensions (Hamann & Schiemann, 2021). Similarly, at the first-order abstraction level, Joo et al. (2011) used various input and output variables in three separate Financial DEA models to measure three-dimensional constructs: total assets efficiency, current assets efficiency, and expense efficiency. At the second-order

[16] This study is limited to first-order and second-order models since they are the most common cases in the Financial DEA literature. Although theoretically, it would be possible for measurement models to exist of various abstraction levels, in this case there are multiple multi-dimensional constructs that form a more abstract construct.

abstraction level, the authors used the three dimensional constructs to represent different aspects of the firm performance phenomenon.

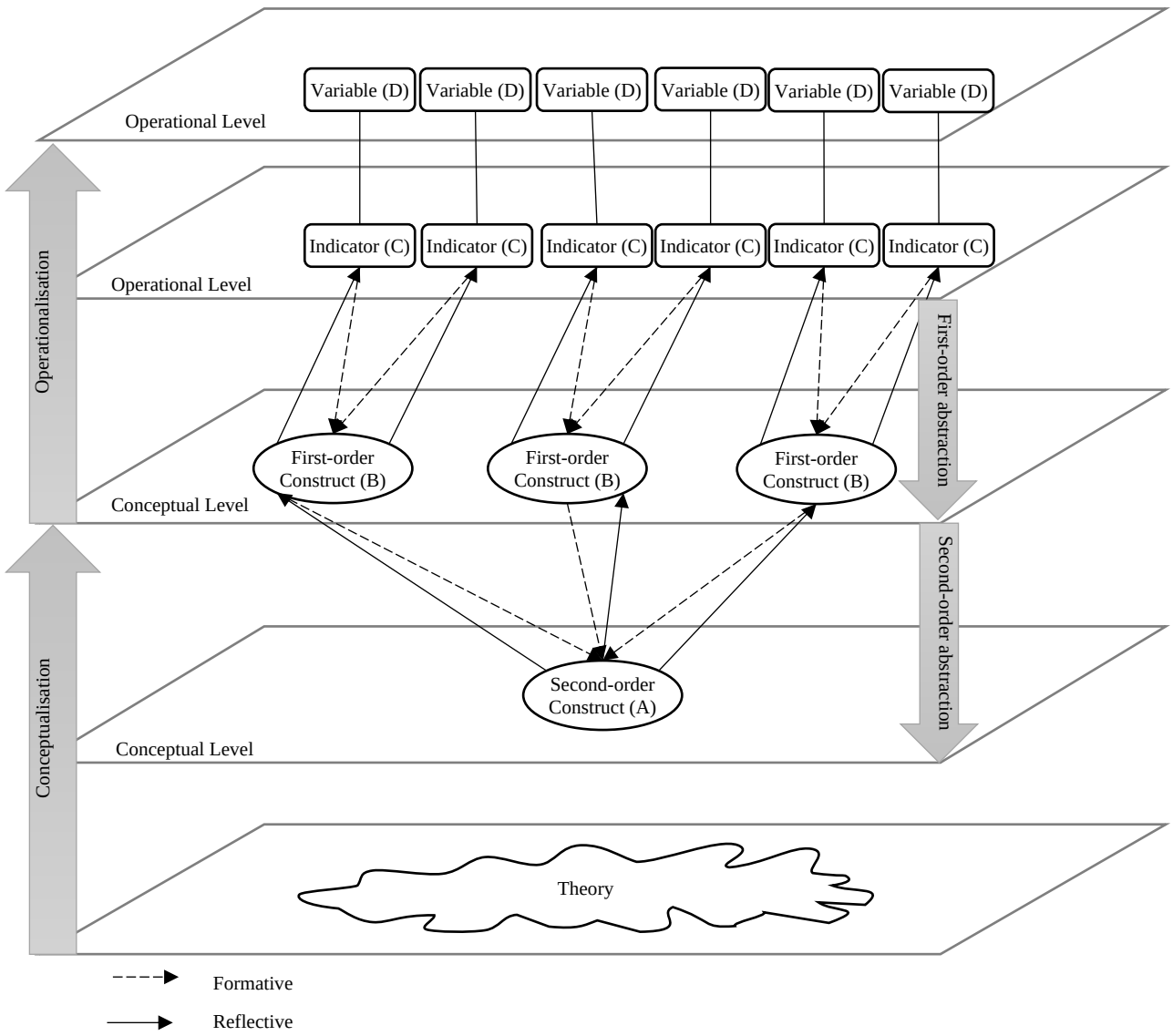


Figure 2 - 7 First-order and Second-order Measurement Models

The higher-order measurement model in this study extends to the second-order measurement modelling, in which the manifestations are not observable. Instead, they are manifested by other indicators (Bollen, 1989; Jarvis et al., 2003). As shown in Figure 2 - 7, the non-observable manifestations (B) are dimensions of second-order constructs (phenomenon)(A), and the dimensional constructs are first-order constructs (Mackenzie et al., 2005).

Dimensions are defined as non-observable manifestations of a construct, and they are essentially sub-constructs of a multi-dimensional construct (a phenomenon) (Bisbe et al., 2007). A multi-dimensional construct (a phenomenon) arises from a holistic concept and connects to several distinct but related dimensions (Figure 2 - 7).

Table 2 - 2 Constructs - Phenomenon Measurement Models (Second Order)

	Reflective	**Formative**
The direction of Causality/ Change	from constructs to dimensions	from dimensions to constructs
Relationship Between Dimensions	correlated, covariant, and interchangeable	not necessarily correlated, covariant, or interchangeable
Result of Dropping Any Dimensions	the domain of phenomenon does not change	the domain of phenomenon changes
Measurement Error	at the dimensional level (δ)	at the construct level (ζ)

This section introduces two second-order models, namely the second-order reflective model and the second-order formative model. It summarises the principle of deciding whether a multi-dimensional construct is reflective or formative Table 2 - 2.

Reflective model

As summarised in Table 2 - 2, dimensional constructs reflect the underlying multi-dimensional construct (the phenomenon) in a second-order reflective model. This model assumes a multi-dimensional construct (the phenomenon) at a deeper and more embedded level of abstraction than its dimensions. The causal relationship flows from the underlying construct (the phenomenon) to the dimensions. The dimensions are the effects of the phenomenon. Thus, a change of the phenomenon leads to a change in the dimensional constructs. Since all dimensions share the same underlying phenomenon, the dimensions are expected to co-vary, and the common area attributes to the underlying phenomenon. Dimensions are correlated and interchangeable. Even dropping a dimension would not change the domain of interest or cause any specification problems (Bisbe et al., 2007; Law et al., 1998).

Formative model

In a second-order formative relationship, a composite construct is a combination of all dimensions. Dimensions are viewed as separate facets and jointly define the characteristics of a phenomenon. The direction of causality is from dimensions to the phenomenon. Therefore, a change in dimensions leads to a change in the phenomenon. Dimensions do not necessarily share the same theme; instead, each dimension forms the phenomenon's definition uniquely. Therefore, omitting a dimension would change the domain of the phenomenon because of the lack of a decisive aspect (Bisbe et al., 2007; Law et al., 1998).

After distinguishing the second-order formative and reflective models, the framework expands to Figure 2 - 7, where there are two levels of abstraction: the first-order abstraction and the second-order abstraction. Both levels of abstraction are empirical abstractions that intend to abstract variables at the operational level towards the theory at the conceptual level. The constructs in the second-order abstraction level (A) are of higher abstraction and closer to the theory than the constructs in the first-order abstraction level (B). The constructs in the first-order abstraction level (B) are dimensional constructs for multi-dimensional constructs (phenomena) (A) in the second-order abstraction level.

Also, in Figure 2 - 7, the arrows represent the direction of causality. The solid arrows indicate the reflective relationships where the constructs cause the indicators, and the indicators reflect the constructs. The dashed arrows indicate the formative relationships where the indicators cause the constructs and indicators form the conceptual constructs.

However, not all applications follow the measurement models precisely, especially at the higher abstraction levels. In practical studies, there may be a central focus of each research bearing multiple constructs. Researchers have the option of choosing general or specific measurement models depending on the researchers' theoretical interests (Mackenzie et al., 2005).

In Financial DEA, for example, Demerjian et al. (2013) were interested in the relationship between managerial quality and the earnings quality constructs. To measure the construct of managerial quality, the construct of firm performance was first measured through a Financial DEA model. The firm performance phenomenon (Figure 2 - 7, A) was not the central focus of the study. The authors treated the firm performance phenomenon as a general construct without specifying the dimensions within it. At the second-order abstraction level, the authors

only used financial efficiency as the dimensional construct (Figure 2 - 7, B) to represent firm performance (Figure 2 - 7, A). At the first-order abstraction level, the financial efficiency was formed by the indicators of economic resource and economic output (Figure 2 - 7, C). These indicators were operationalised to seven accounting variables as inputs and one accounting variable as the output in the Financial DEA model (Figure 2 - 7, D).

However, when the dimensional constructs measured by Financial DEA are the central focus of research, multiple dimensional constructs can be found (Figure 2 - 7, B). For example, Joo et al. (2011) specifically distinguished three-dimensional constructs within firm performance phenomenon: total asset efficiency, current asset efficiency, and expense efficiency (Figure 2 - 7, B). At the second-order abstraction level, these three-dimensional constructs, representing separate aspects of firm performance, formed the firm performance phenomenon (Figure 2 - 7, A). At the first-order abstraction level, these three-dimensional constructs were represented by various indicators (Figure 2 - 7, C). The inputs of total asset efficiency were represented by the indicators of both current and non-current asset categories. The inputs of current asset efficiency were represented by current asset categories only. The inputs of expense efficiency were represented by expenses categories only. These indicators were ultimately operationalised into corresponding accounting variables (Figure 2 - 7, D).

2.6. Chapter Summary

This chapter has reviewed literature supporting the definition and modelling of Financial DEA. As noted, Financial DEA is defined as a DEA model that solely incorporates accounting data from financial reports. The definition of Financial DEA leads to several inter-related fields.

Data Envelopment Analysis explains the modelling features of Financial DEA. The methodological issues of DEA may be relevant to Financial DEA modelling. The duality theory and DEA axioms may apply to the application of Financial DEA. Accounting information introduces additional information to Financial DEA compared to the physical measures in conventional DEA. Firm performance introduces alternative views and meanings to the research context of Financial DEA. Measurement models introduce various views of the modelling process of Financial DEA. These fields comprise the domain of Financial DEA.

By reviewing the relative literature, several gaps were identified. First, a comprehensive typology is needed to understand how Financial DEA models measure firm performance. An overview is needed to cover the dimensions of firm performance measured by Financial DEA, and how they have been measured. Second, methodological issues in the Financial DEA application need to be summarised, and empirical tests are needed to quantify the impact of potential methodological issues.

The next chapter will outline the research approach of this study to fill the gaps identified in this chapter. Specifically, an analytical approach will be carried out, including a typology of Financial DEA studies, a conceptual framework of Financial DEA and simulated and empirical tests need to be carried out. These methods and methodologies will be discussed in the next chapter.

Chapter 3: Research Approach

3.1. Chapter Introduction

This chapter describes the research approach of this study. Section 3.2 reviews the methodology and the common philosophical views held in accounting. This study applies a positivist view, and the methods are designed accordingly. Section 3.3 introduces the methods used in this study. This study carries out a two-phase analytical approach. Phase I (Chapter 4) builds a Financial DEA literature typology and identifies potential sources of Financial DEA measurement errors at the modelling level. Phase II (Chapter 5) quantitatively examines selective methodological issues identified in Phase I, based on the literature typology and the conceptual framework at the modelling level. Section 3.4 summarises the chapter.

3.2. Methodology and Paradigm in Accounting

Methodology is the study of method, and it covers the procedures and goals of a particular discipline, and enquiry into the way in which that discipline is organised (Creamer, 2018; Flew, 2002; Ghauri & Grønhaug, 2010). As illustrated in Figure 3 - 1, methodology is the understanding of how methods are constructed or how an operative paradigm is developed. An operative paradigm relates a methodological view to a specific study area. Methodological views are different schools of philosophy that are determined by various ultimate presumptions made. The ultimate assumptions lead to various views of the problem, which include multiple elements, such as ontology (beliefs about knowledge) and epistemology (beliefs about reality) (Arbnor & Bjerke, 2009).

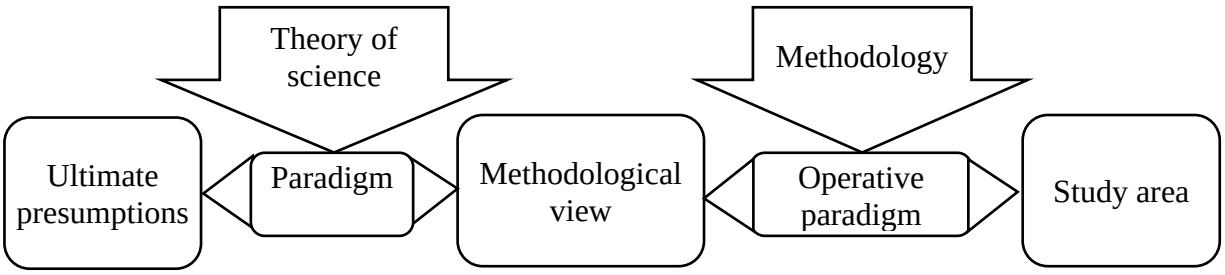


Figure 3 - 1 Methodology of Researching Knowledge
Adapted from Arbnor and Bjerke (2009)

However, the assumption elements are not immutable or exhaustive but historically specific (Chua, 1986). With assumptions, the way of conceptualising reality is defined as paradigms. Paradigm is the theory of science that links the ultimate assumptions and methodological views. Paradigm choices were initially defined by Kuhn (1962) as value-based decisions between incompatible models of scientific life. Since then, there has been a range of proposals to classify paradigms. This section focuses on the accounting view of paradigms (Chua, 1986).

Hopper and Powell (1985) based on earlier sociological approaches (Burrell & Morgan, 1979), classified management accounting literature according to two sets of assumptions: (a) about social science and (b) about society. First, the assumptions about social science related to the subjective-objective view. The subjective end was known as “German idealism”, which assumed that reality lay in ideas from a nominalism view and took anti-positivist and voluntarist views in epistemology, human-nature assumptions, and ideographic methodological approaches. The objective end is described as “sociological positivism”, which attempted to apply methods derived from natural science to sociological studies (Burrell & Morgan, 1979).

Second, the assumptions about society are either orderly change or fundamental conflict, that is, radical change (Burrell & Morgan, 1979). However, Hopper and Powell (1985) argued that while sociological paradigms were useful for indicating the nature and range of alternative approaches, it would be wrong to claim that all accounting research can be neatly classified into these categories. Later, Chua (1986), based on the work of Burrell and Morgan (1979), and Hopper and Powell (1985), classified accounting research into three types based on different paradigms and world views.

These three classifications are based on differing world views, including: (a) beliefs about knowledge, (b) beliefs about physical and social reality, and (c) the relationship between theory and practice (Chua, 1986).

First, the beliefs about knowledge cover epistemological and methodological assumptions.[17] Second, the beliefs about physical and social reality cover ontology, human intention, and social order (Chua, 1986).[18] Last, the relationship between theory and reality assumes that society needs to be changed by theories. A low level of change view is satisfied with the current order in society and takes the existing institutional structures for granted. Researchers use theories to understand the social order and explain how society functions. By contrast, a high level of change focuses on conflicts in society, and theories argue that society needs to be changed (Chua, 1986; Hopper & Powell, 1985).

As a result, Chua (1986) identified three types of accounting research, being: (a) the mainstream accounting view, (b) the interpretive view, and (c) the critical view (Chua, 1986) (Figure 3 - 2). First, mainstream accounting takes a positivist view and assumes the reality is epistemologically verifiable or falsifiable by observations. It methodologically follows scientific rigour and is generalisable, and ontology is objective and independent. Regarding human intention, mainstream accounting research also assumes humans intend to maximise the utility of the firm and its individuals. In terms of societal order, this stream assumes that both society and organisations are stable. The purpose of knowledge to practice is that accounting serves as a means but not an end. Accountants should only provide accounting information to decision-makers but not influence any decisions (Chua, 1986).

In management accounting, mainstream research started to flourish in the 1960s with a strong influence from the neoclassical economy. In the beginning, management accounting was

[17] Epistemology assumes the nature of knowledge (Blackburn, 2016). Burrell and Morgan (1979) assumed the way to gain knowledge ranged from objective positivism to subjective anti-positivism. Positivists take the objective view and seek to explain and predict the social world by searching for regularities and causal relationships. They believe the world has an objective existence and can only be accessed by observation. Contrarily, anti-positivists hold that the world is relativistic and can only be understood from individuals involved in the activities (Burrell & Morgan, 1979; Ryan et al., 2002).

[18] Ontology assumes the nature of existence (Flew, 2002). Especially in social science, ontology is about whether the "reality" to be investigated is external to the individual or the product of individual consciousness. Burrell and Mongan (1979) assumed that reality ranged from the objective end -- realism to the subjective end -- nominalism. On the objective end, realists hold that reality exists within the objective, concrete construction, and the social world is external to individual cognition. The social world is a real-world made up of hard, tangible, and relatively immutable structures. On the subjective end, nominalists state that reality exists within the subjective mind, and the social world external to individual cognitions is made up of nothing more than names (Burrell & Morgan, 1979; Ryan et al., 2002).

normative, with prescriptions for what ought to happen. In the 1970s and 1980s, there was an emphasis on a positivist (economic) view to explain observed mainstream research, and this influenced accounting practices (Ryan et al., 2002).

Second, the interpretive paradigm assumes that epistemological knowledge is gained from subjective interpretation and specific explanation. Methodologically, participation and case studies are encouraged. Ontologically, reality is considered subjectively created. Human actions have intentions, and there is a certain social order. The relationship between theory and practice explains the action, especially to understand how the social order is produced (Chua, 1986). Similar to positivism, researchers assume a low level of change in society, and interpretive researchers explain how society holds together (Chua, 1986). The interpretive view assumes a low level of theorisation, which means that theories are difficult to generalise. Therefore researchers are directly involved in research and use their perceptual skills rather than following theoretical processes (Laughlin, 1995).

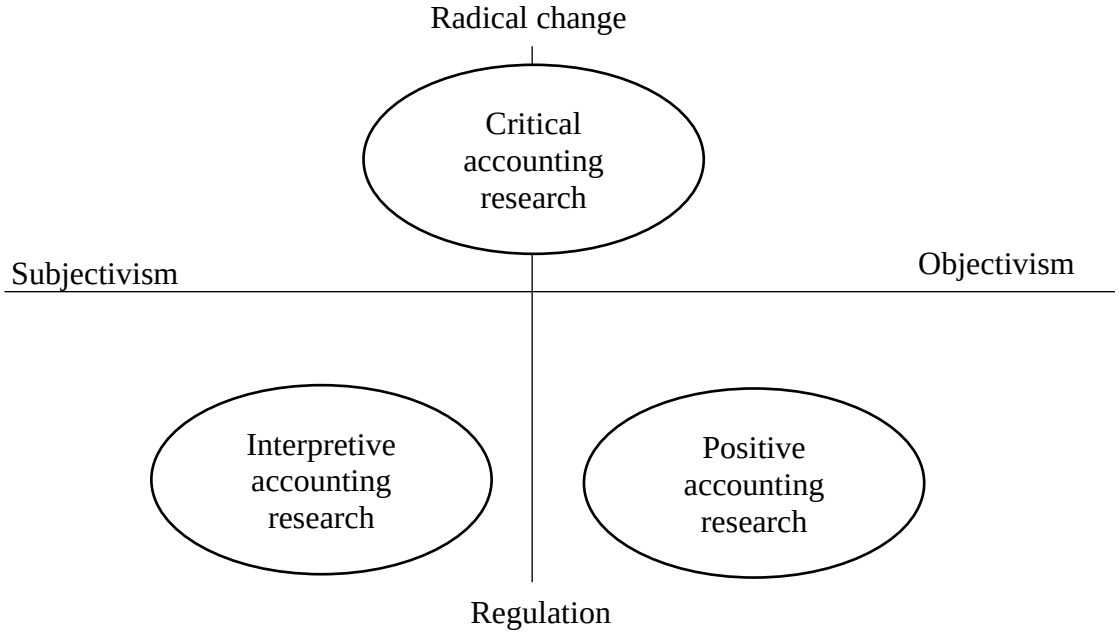


Figure 3 - 2 Typology of Accounting Research
Adapted from Ryan et al. (2002)

Third, the critical paradigm assumes that regarding the nature of knowledge (epistemology), theories are temporal and context-bonded. Moreover, methodologically, case studies are usually employed. As for social reality, ontological research objects can be understood by studying the historical development and change of relationships between individuals. Human beings reproduce objective reality through subjective interpretation. Society is assumed to have fundamental conflicts due to injustice in social and political domains. The role of theory has a critical perspective that identifies and removes current practices (Chua, 1986). Under this critical view, social structure is a result of social action. Meanwhile, social action is also determined by social structure (Hopwood, 1987; Miller & O'Leary, 1987; Ryan et al., 2002).

This study assumes the positivist paradigm and believes that the Financial DEA model is a tool to gain objective knowledge of firm performance. Specifically, this study assumes that social reality, especially firm performance, is objective, and theory is considered true if it is not falsified by empirical evidence. Based on the positivist methodological view, research methods will be built and presented in the next section.

3.3. Research Methods

3.3.1. Analytical Approach: Induction and Deduction

The analytical approach is based on an analytical and methodological view, and the positivist paradigm. With the assumption of this positivist paradigm, the analytical approach assumes that the nature of reality is objective and independent. This approach also assumes that knowledge can only be obtained in a rigorous scientific way as a picture of business reality. Under the positivist assumptions, a theory is operationally tested by logic and mathematics in verifying propositions and follows formal rules from natural science. The research results are refined and logical models, and generalisable cases (Arbnor & Bjerke, 2009).

An analytical approach has a cyclical feature that starts with the facts of one cycle and ends with facts beginning the next cycle through inductive and deductive research. The first step – induction draws a general conclusion from empirical observations. The second step – deduction is a logical analysis of general theory through empirical scrutiny (Arbnor & Bjerke, 2009).

In this study, Phase I is inductive. Theoretically, it investigates whether Financial DEA applications fit with the production theory and the benchmarking approach identified in the DEA literature (section 2.2.3). The output (Chapter 4) is a Financial DEA literature typology and Financial DEA frameworks at the construct level and at the modelling level. These frameworks provide a theoretical starting point for Phase II of the study. Phase II is deductive (Chapter 5), where both simulated data and archival data will be used to quantitatively examine the impact of different model choices. The results of both phases will be synthesised to develop a conceptual foundation for Financial DEA (Chapter 6).

3.3.2. Phase I: Conceptualisation

Phase I aims to develop a conceptual framework to provide a conceptual foundation for Phase II. An inductive approach was chosen in Phase I since the literature does not provide a comprehensive view of the conceptual foundation of Financial DEA. The research process started with collecting the literature. The first output of Phase I is a conceptual typology, which is a conceptual classification, within which the categories represent the type of concepts rather than empirical cases (Bailey, 1994). Unlike typology, taxonomy provides subsequent identification of empirical cases for conceptual typologies (Bailey, 1994). The typology was developed by interactively classifying the Financial DEA models from the literature by the types of inputs and outputs used. As a result, the typology provides Financial DEA research themes, domains, and dimensions (Patton, 2002).

The second output of Phase I is a conceptual framework at the modelling level identifying potential methodological issues in the Financial DEA application, which is the basis of the analysis in Phase II. The framework provides a detailed plan for Phase II deductive research by pointing out potential methodological issues in the Financial DEA application (Ravitch & Riggan, 2012).

3.3.2.1. Data collection: secondary qualitative data

For the Financial DEA literature typology, the literature review was restricted to Financial DEA empirical studies, which are defined as those that exclusively incorporate financial accounting data in DEA models. There are two steps in the literature collection process. First, literature is collected by searching for keywords in the Financial DEA definition, including different combinations of “DEA”, “Data Envelopment Analysis”, “Financial Report Analysis”, “Financial”, and “Accounting” in Google Scholar and Scopus. Each identified

paper is later screened to confirm that all inputs and outputs are exclusively financial accounting variables in the DEA models. Since Financial DEA is an emerging stream of research, the sampling process is opened to all academic journals without constraining the search for journal quality or journal area.[19]

Second, studies are collected using the snowball sampling method to search for new studies based on the citations in the studies identified. The snowballing sampling method is a multistage method starting with an initial sample where a given set of criteria are applied repetitively (Krippendorff, 2013).

For the methodological issues in Financial DEA, the literature of methodological issues in DEA application (section 2.2.4), and the literature of measurement models (section 2.5) are synthesised with the Financial DEA studies identified from the typology. The accounting information in Financial DEA, which is different from the physical measures in the conventional DEA, is impacted by business activities, including pricing strategy, operational activity, and accounting regulations. Also, Financial DEA models are mostly used to measure firm performance, and the literature of measurement models can help identify issues during the modelling process.

3.3.2.2. Systematic analysis

The literature collected will be analysed to assist the development of a conceptual framework for Financial DEA. For the Financial DEA literature typology, a qualitative analytical process is applied to systematically identify and interpret patterns of meaning (themes) within the literature (Clarke & Braun, 2017). The analytical process aims to produce a typology of Financial DEA, a classification of Financial DEA research themes, and a conceptual framework of measurement models in Financial DEA.

The qualitative analytical process contains four fundamental operations, including: (a) categorisation**,** (b) abstraction, (c) comparison and (d) integration, and two tactics – (e) iteration and (f) refutation (Ghauri & Grønhaug, 2010; Spiggle, 1994). The six operations are neither mutually exclusive nor in sequential order. This is especially as the tactics suggest, researchers need to move back and forth between steps (iteration) and sometimes disconfirm

[19] Considering the novelty of "Financial DEA", conference proceedings, working papers, and book chapters were also collected in the sample set. The journal quality and journal area were coded and are analysed in Chapter 4 (section 4.2.1).

the emerging analysis with negative cases (refutation). The Financial DEA literature will be analysed by DEA models. For example, if one study has two Financial DEA models, each model is treated separately. For each model, the features (e.g. input and output variables) were recorded and coded into a spreadsheet manually (Appendix 1).

In this study, a spreadsheet was created, where rows were in the unit of Financial DEA models and columns represented different variable categories. The categories are based on the two views of DEA discussed in section 2.2.3, and these are the productive efficiency view and the benchmarking view. Under the productive efficiency view, productive elements (i.e. material, labour, capital, and output volume) are used to classify inputs and outputs. For example, "the cost of machine" is one accounting variable in the "capital" category. Under the benchmarking view, accounting variables from the accounting equation (i.e. *Assets + Expenses = Liabilities + Equity + Revenue*[20]) are used to classify inputs and outputs. For example, "operating expense" is one accounting variable in the "expenses" category.

For the framework at the modelling level, the generic methodological issues in DEA and measurement models were analysed against the Financial DEA literature. Common issues in Financial DEA are summarised with illustrative examples identified from the Financial DEA literature typology.

In summary, Phase I conceptualisation uses an induction research approach to first develop a typology of the Financial DEA literature, categorising the research domain and the themes of the Financial DEA studies. Moreover, based on the typology, the generic methodologic literature of DEA (section 2.2.4), and measurement models (section 2.5), this phase identifies potential sources of methodological issues in the Financial DEA application. The next section introduces the methods for Phase II, which deductively tests the methodological issues identified in Phase I.

3.3.3. Phase II: Empirical Examination

The purpose of Phase II is to empirically and selectively examine the conceptual framework developed in Phase I. The framework developed in Phase I will identify different measurement models and potential sources of measurement errors. The conceptual

[20] Note this is the expansion of the general accounting equation: Assets = Liabilities + Equity, which separately identifies the current period movements in equity of revenue and expenses.

framework at the modelling level will be used to develop selective propositions. In Phase II, the propositions are operationalised to empirical tests to examine the extent of the impact of different sources of measurement error on the results of Financial DEA. Note that the empirical test in Phase II does not cover all measurement errors identified in Phase I. Only the key sources of measurement errors of particular importance to Financial DEA based on the literature review and literature typology will be examined.

Empirical tests are advocated by empiricists who believe that knowledge is based on experience (either qualitative or quantitative data), as opposed to purely logical relations (rationalism) (Goodwin & Goodwin, 2017; Ryan et al., 2002). The empirical data are quantitative using both simulated data and archival data collected from financial databases. This quantitative empirical data will be used to examine the impact of measurement errors identified in the conceptual framework.

3.3.3.1. Simulations with artificial data

Simulated data or artificial data are used in simulations. Compared with empirical tests with archival data, simulations have the advantage of controlling for extraneous factors; therefore, comparisons between scenarios can be made (Hatami-Marbini & Toloo, 2017). Sometimes the scenario compared against is assumed to be the "truth", where "true efficiency" is generated (Hatami-Marbini & Toloo, 2017). By comparing across scenarios, the impact on DEA models can be evaluated. When comparing scenarios, assumptions of the predetermined production functions need to be made (Harrison et al., 2012). However, simulations have a weakness in that the predetermined function needs various assumptions, such as the parametric function and distribution properties. As a remedy, archival data are also incorporated in the Phase II tests to examine whether the relationships found in the simulated test are realistic.

Various assumptions need to be made to build simulations. First, for predetermined functions, there are two views, namely that from the economics perspective and that from the operational research perspective. The economics view suggests that the predetermined functions are of a production nature, such as the Cobb-Douglas function (Banker et al., 1993; Oh & Shin, 2015; Ruggiero, 2007) and the transcendental logarithm function (Andor & Hesse, 2014; Cordero et al., 2009). This study used the Cobb-Douglas function as the predetermined production function since it is relatively widely adopted (Harrison et al., 2012;

Perelman & Santín, 2009; Ruggiero, 1999). Alternatively, the operational research view advocates that the predetermined functions do not need to be a production function since DEA is a nonparametric function. Instead, randomised DMUs can form an efficient frontier, and inefficiencies can be applied later (Khezrimotlagh et al., 2019). This study took different predetermined functions in individual tests based on the different assumptions.

Second, for the parameter distributions, both normal distribution (Harrison et al., 2012; Ruggiero, 2007) and uniform distribution (Andor & Hesse, 2014; Ruggiero, 1998) are commonly used by researchers. This study chose normal distribution since this study focuses on the application aspect of DEA, and the business reality is assumed to be closer to a normal distribution according to the central limit theorem (Fischer, 2010).

3.3.3.2. Empirical tests with archival data

Empirical tests with archival data are incorporated in this study since, in the simulations, the predetermined parameters may not fully capture the real-world features. As a remedy, empirical tests with archival data are carried out.

The archival data are retrieved from the COMPUSTAT North American database, covering five years of financial data. The five-year timeline provides a large enough sample for analysis while reducing the likelihood that significant technological change had occurred during the test period (Färe et al., 1994; Golany & Roll, 1989).

The sample includes six industries as sub-samples, based on the Fama-French 48 industry classification. These industries are selected to provide a representative selection of industries (Demerjian et al., 2012; Fama & French, 1997). The six industries are the automobile industry (SIC 2296 - 3799), the box industry (SIC 2440 - 3412), the clothing industry (SIC 2300 - 3965), the food industry (SIC 2000-2099), the gold industry (SIC 1040 - 1049), and the personal services industry (SIC 7020 - 8899).

The six industries are selected for the following reasons. First, considering DEA is sensitive to heterogeneous factors, the industries are selected, according to Demerjian et al. (2018), where the mean efficiency scores are above 0.7, indicating potentially high levels of homogeneity (Dyson et al., 2001). The gold industry is an exception, with a mean efficiency score of 0.320 (Demerjian, 2018). However, this industry is included to provide an example of less homogeneous data. Also, the homogeneity of the industries is reviewed by examining

industry descriptive features. For example, the personal services and food industries are relatively more heterogeneous since the products and services covered vary significantly from the other selected industries. Therefore, the business activities in these industries are relatively heterogeneous, with various inputs and outputs. In addition, the gold industry is also relatively heterogeneous since the operating environment is highly subject to factors such as the mining locations.

Second, the six industries are chosen to provide a diversity of industrial features. For instance, the clothing and personal services industries are relatively labour intensive to provide sales and services. By comparison, the food and gold industries are relatively capital intensive since they rely on production machinery. Moreover, industries are selected to represent varying firm sizes. For example, the DMUs from the automobile industry tend to be relatively large, considering the machinery and scope requirements. In contrast, DMUs from the personal services industry such as personal laundry and beauty shops are more likely to be small and medium-sized businesses due to the nature of their business.

Third, the sample size is considered since the relative relationship between the sample size and the model size influences the discriminatory power (Dyson et al., 2001) (section 2.2.4). Relatively large industries such as the automobile industry (initial sample size = 388), the food industry (initial sample size = 360), the gold industry (initial sample size = 409) provide a comparison with relatively small industries such as the box industry (initial sample size = 59) and the personal services industry (initial sample size = 55). The final sample sizes are different from the initial sample sizes after the data cleaning process. The data cleaning process first deletes DMUs with missing values and zero values based on the input and output variables required. This leads to different final sample sizes in Test Two (section 5.3.2.1) and Test Three (section 5.4.1.1) since the financial models are designed with different accounting variables.

Next, all variable values are inflated to the 2019 financial year using the consumer price index (CPI) (Rouse & Tripe, 2016). Last, for models incorporating negative values, the negative values are translated to positive values by adding a positive constant value to each DMU so that any negative values are translated into positive ones in the DEA algorithm (Bowlin, 1999; Charnes et al., 1983; Seiford & Zhu, 2002). Alternative treatment of deleting negative values is also carried out to check the robustness of the results.

3.3.3.3. Model performance criteria

To assess the performance of the Financial DEA models, or the level of impact on the Financial DEA results, reliability and validity are used as the standards for positive research (Edwards, 2003; Ghauri & Grønhaug, 2010).

Reliability is indicated when the measures sharing the same method show convergence. When the methods are different, the convergence of measures indicates convergent validity, and the divergence of measures indicates discriminant validity. Both convergent validity and discriminant validity are subcategories of trait validity, which focuses on the relationship between constructs and measures. This relates to the measurement models discussed in section 2.5, explaining the modelling process between the Financial DEA models and the firm performance phenomenon.

Trait validity and nomological validity are two subcategories of construct validity (Edwards, 2003). Construct validity explains the relationship between constructs and quantifies the degree to which a measure captures its theoretical construct (Cronbach & Meehl, 1955; Edwards, 2003; Nunnally & Bernstein, 1994).

This study focuses on trait validity since it explains the relationship between the construct and measures (section 2.5). The nomological validity focuses on the relationship between measures of a construct with measures of other constructs under relevant theory (Bisbe et al., 2007; Cronbach & Meehl, 1955; Edwards, 2003). Nomological validity is not the focus of this study and, therefore, will not be discussed further. Similarly, there are other classifications of validity, which are not relevant to the relationships between constructs and measures. Therefore, they will not be discussed further in this section either. For example, content validity explains the degree to which a measure represents a particular domain of content (Edwards, 2003; Nunnally & Bernstein, 1994). Content validity is not assessed using empirical or statistical procedures, but instead relies on “appeals to reason” (Nunnally & Bernstein, 1994). Also, criterion-oriented validity explains the relationships between the measure of interest and the criterion measures (Nunnally & Bernstein, 1994). This is also not the focus of this study and therefore, it will not be discussed further.

In this study, different Financial DEA models represent different methods. Therefore, when comparing across Financial DEA models, the performance is evaluated against the convergent validity and discriminant validity within trait validity and construct validity.

The convergence of measures using different methods represents convergent validity. Divergence among measures using the same or different methods demonstrates discriminant validity (Campbell, 1960; Campbell & Fiske, 1959; Edwards, 2003). There is a lack of an agreed method in the literature to test the convergent and discriminant validity (Jarvis et al., 2003; Petter et al., 2007). For instance, researchers suggested that the development of structural equation modelling (SEM) is one method to assess construct validity (Jarvis et al., 2003). Also, validation models such as the principal components model and the common factor model are used to determine whether the measures of the same (different) constructs cluster (separate) (Edwards, 2003; Harman, 1976; Kim, 1978).

A systematic approach named multitrait-multimethod (MTMM) matrix was proposed by Campbell and Fiske (1959). The MTMM matrix calculated correlations among several measures of constructs using different methods (Edwards, 2003). The MTMM matrix also provides criteria to assess convergent and discriminant validity. This study adapted selective correlations and criteria suggested by the MTMM matrix (Campbell & Fiske, 1959; Edwards, 2003). And adapted them with DEA model evaluation criteria to investigate the impact on convergent and discriminant validity of the Financial DEA results.

This study uses three criteria to assess Financial DEA model performance, being: (a) the Pearson correlation, (b) the Spearman's ranking correlation and (c) the mean absolute deviations, MAD. These criteria are selected based on the purpose of the study and the common criteria used in recent DEA literature.

This study used (a) the Pearson correlation, and (b) the Spearman's ranking correlation between the Financial DEA results and the comparison point to evaluate the Financial DEA model performance. Correlation is the method suggested by researchers to assess convergent and discriminant validity (Campbell & Fiske, 1959; Edwards, 2003) and by researchers to assess DEA model performance. Recent DEA research has used three types of correlation methods: the Pearson correlation, Spearman's ranking correlation, and the Kendall correlation. Spearman's ranking correlation is the most common method used in DEA (Andor & Hesse, 2014; Oh & Shin, 2015). It is a nonparametric test that measures the relative change in DMU rankings across scenarios (Giorgio et al., 2016; Ruggiero, 2004). DEA researchers also widely use the Pearson correlation (Hatami-Marbini & Toloo, 2017; Ruggiero, 2007). It measures the linear relationship between the efficiency scores. By comparison, the Kendall

correlation is comparatively less used (Zelenyuk, 2020). It is a nonparametric test that measures the strength of dependence between two variables (Kendall, 1938).

Further, the former two correlation methods are consistent with the Financial DEA research purpose. Most Financial DEA researchers aim to measure firm performance relative to its peers, and they focus on the relative efficiency scores (Demerjian et al., 2012; Demerjian et al., 2013) and ranking (Mahajan et al., 2014) to interpret firm performance.

Convergent validity is defined as the convergence of measures of the same construct. Statistically, it is evidenced as correlation values that are large and significantly different from zero (Campbell, 1960; Campbell & Fiske, 1959; Edwards, 2003). In this test, any correlation coefficient greater than 0.8 with a p-value less than 0.1 is considered evidence of convergent validity. Discriminant validity is defined as the divergence among measures of different constructs. Statistically, it is evidenced as the correlation values being effectively smaller than the convergent values (Campbell, 1960; Campbell & Fiske, 1959; Edwards, 2003). In this test, any correlation coefficient of less than 0.2 or a p-value greater than 0.1 is considered evidence of discriminant validity.

This study used (c) the mean absolute deviations, MAD, to calculate the difference between the estimated efficiencies and the comparison point following Banker et al. (Andor & Hesse, 2014; Banker et al., 1993; Santín & Sicilia, 2017; Zelenyuk, 2020). The calculation method of MAD is mathematically expressed as below:

$$MAD = \frac{1}{NR}\sum_{r=1}^{R}\sum_{j=1}^{N}|\hat{\theta}_{rj} - \theta_{rj}|$$

where N is the number of DMU, R is the number of iterations, $\hat{\theta}$ denotes the estimated efficiency, and θ represents the efficiency of the comparison point (Andor & Hesse, 2014; Santín & Sicilia, 2017).

A low MAD indicates that the estimated efficiency scores are close to the efficiency score of the comparison point. Banker et al. (1993) found that low measurement errors (e.g. measurement error variance equal to 0.002 or 0.004) lead to the DEA results being dominated by the inefficiency distribution, and MAD decreases with the increase in sample size. By comparison, high measurement errors (e.g. measurement error variance equal to 0.02 or 0.04)

lead to the DEA results being dominated by the measurement errors, and MAD increases with sample size.

The MAD method has been widely used by researchers (Giorgio et al., 2016; Oh & Shin, 2015; Zelenyuk, 2020). More importantly, the MAD method provides additional information to the correlations of the individual DMUs' distance from the Financial DEA estimation and the comparison point (Andor & Hesse, 2011).

3.4. Chapter Summary

This chapter covered the research approach of this study, including the paradigm, the analytical approach, and the research design.

This study is built on the positivist paradigm, assuming that reality and knowledge are objective, and researchers explain the relationships in society. Specifically, Financial DEA is a means to measure firm performance, and firm performance is an objective phenomenon. The study follows an analytical approach in an inductive-deductive two-phase approach. Phase I inductively builds a conceptual framework based on the qualitative evidence from the Financial DEA literature. The framework identifies the theme and domain of Financial DEA applications. As a result, different conceptual measurement models of Financial DEA are identified, along with potential sources of measurement errors. Phase II examines the impact of the sources of measurement errors specific to Financial DEA using simulated and archival data.

Chapter 4 and Chapter 5 contain the findings from Phase I and II, respectively. Chapter 4 reports the descriptive typology of Financial DEA research and summarises the themes and domains of application. The typology results are used to develop a framework at the modelling level of Financial DEA research and potential sources of measurement errors. Chapter 5 investigates the measurement errors identified by the conceptual framework at the modelling level and the impact of accounting variable choice on Financial DEA modelling. Chapter 6 synthesises the findings of the two-phase study and the implications for Financial DEA researchers.

Chapter 4: Phase I – Conceptualisation

4.1. Chapter Introduction

This chapter reports the result of the analysis of the Financial DEA literature. It aims to answer the first and second research questions. It based on the first two research questions to develop a range of propositions for the third research question.

This chapter is structured as follows. Section 4.2 aims to answer the first research question: *What dimensions of firm performance do Financial DEA models measure, and how have they been used*. To answer this question, this section provides a typology of Financial DEA models, including the classification of dimensional constructs of firm performance, the accounting indicators, and accounting variables used in Financial DEA. Section 4.3 aims to answer the second research question: *What are the methodological issues when applying Financial DEA?* To answer this question, this section discusses the potential methodological issues that could arise in the Financial DEA application. Section 4.4, aims to develop the propositions to be tested to answer the third research question: *What are the empirical impacts of methodological choices on the results of Financial DEA?* To answer this question, this section, based on the Financial DEA typology and selective methodological issues, develops propositions for the empirical tests in Chapter 5. Section 4.5 summarises the chapter.

4.2. Typology of Financial DEA

Following the inductive approach, this section reports the patterns of Financial DEA literature in the form of a typology. This typology includes the classification of dimensional constructs of firm performance, the accounting indicators, and accounting variables used in Financial DEA studies. Section 4.2.1 describes the scope of the construct of firm performance measured by Financial DEA. Next, section 4.2.2 describes the categories of dimensional constructs of firm performance measured by Financial DEA. Finally, section 4.2.3 describes the categories of the indicators and variables used in the Financial DEA literature.

4.2.1. **Descriptive Analysis of Financial DEA**

In total, there were 210 papers identified in the literature search that incorporate 248 Financial DEA models. The typology was developed by classifying the Financial DEA models from the literature by the inputs and outputs used (Appendix 1).

As discussed, Financial DEA studies have developed since 1990 (Smith, 1990), with the majority published in the past 15 years. There were 210 models identified and 248 models in total. The breakdown of the number of models per year is reported in Table 4 - 1 below.[21]

Table 4 - 1 Number of Financial DEA Models

Year	Count of models	Year	Count of models
1990	3	2008	19
1992	1	2009	16
1995	1	2010	8
1996	2	2011	14
1997	3	2012	7
1998	3	2013	11
1999	4	2014	21
2000	3	2015	11
2001	2	2016	13
2002	7	2017	21
2003	4	2018	24
2004	2	2019	13
2005	3	2020	21
2006	4	2021	1
2007	6	***Total***	***248***

[21] Years are not listed if there were no Financial DEA papers identified from those years.

As illustrated in Table 4 - 2 below, the quality of the publications is relatively high, with over half (55.87%) of the models published in A or A* journals according to the Australian Business Deans' Council (ABDC) journal quality list.

Table 4 - 2 Quality of Financial DEA Studies

ABDC ranking	Count of models
A*	60
A	78
B	61
C	25
Book	2
NA	22
Total	***248***

The field of research covers economics, operational research, management accounting, financial accounting, and finance. The literature's key themes relate to firm performance measurement, such as financial performance, operational performance, and productivity. The next section (section 4.2.2) will discuss the detailed classification of the dimensions of constructs covered by the 210 Financial DEA studies.

4.2.2. Categories of Dimensional Constructs of Firm Performance

This section describes the dimensional constructs of firm performance measured by Financial DEA models identified in the literature. As discussed in section 2.2.3, DEA models can provide alternative views of firm performance, being either productive efficiency (production frontier) or benchmarking (best-practice frontier) (Cook et al., 2014). As illustrated by Figure 4 - 1, the typology starts with the two views proposed by Cook et al. (2014). When DEA is used to measure productive efficiency, variables proxy the productive elements in a physical production process and construct a productive frontier. Alternatively, when DEA models are used to benchmark performance, variables are performance attributes to maximise (outputs) or minimise (inputs) and construct a best-practice frontier (Cook et al., 2014).

The productive frontier category follows the reflective measure modelling discussed in section 2.5.2.1. When measuring productive efficiency, accounting variables in Financial DEA closely proxy the productive elements in the underlying physical production process. The best-practice frontier category follows the formative measurement modelling discussed

in section 2.5.2.2. When benchmarking performance, accounting variables in Financial DEA are the inputs being the performance attributes to minimise, and the outputs being the performance attributes to maximise.

Section 4.2.2.1 describes the dimensional constructs within the productive frontier category. Section 4.2.2.2 describes the dimensional constructs within the best-practice frontier category. Finally, section 4.2.4 summarises the breadth of dimensional constructs identified and links them to higher-order measurement models (firm performance phenomenon).

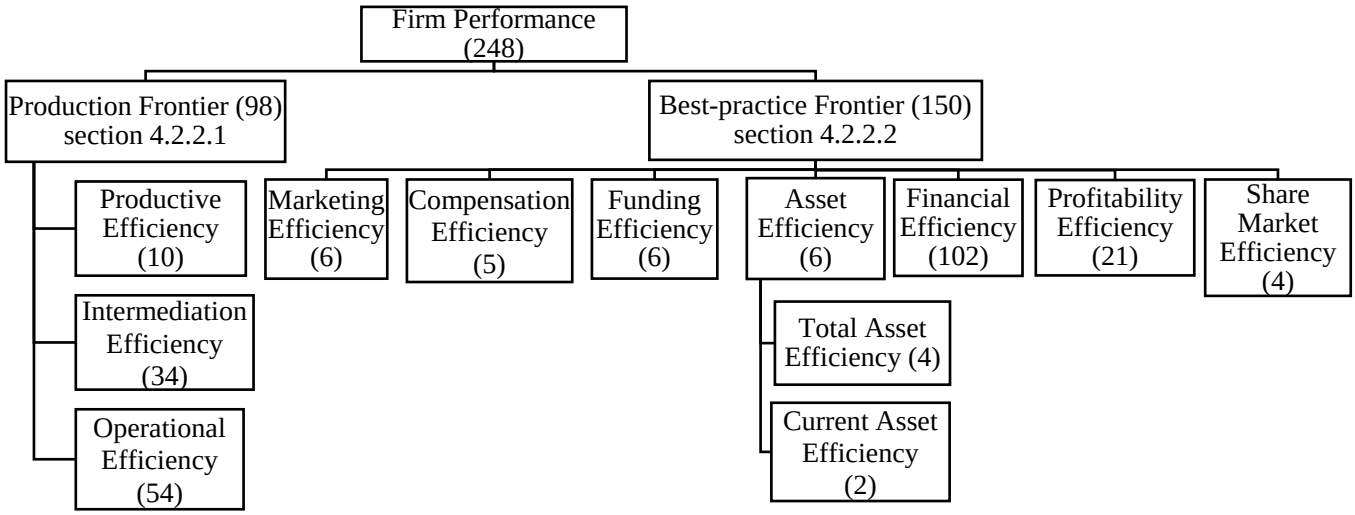


Figure 4 - 1 Categories of the construct of firm performance (numbers of models identified)

4.2.2.1. Productive efficiency

Financial DEA models closely reflect an underlying production process within the productive efficiency category and construct a production frontier. According to economics, the production process primarily covers production elements, including labour, material, and capital. Therefore, these three production factors and the output, production volume are the indicators that signal productive efficiency.

In conventional DEA, the factors of production are operationalised to physical measures due to the close relationship with the physical production process. However, due to data accessibility, Financial DEA research uses accounting variables to proxy physical production factors. For example, Saranga (2009) explained that financial variables in her Financial DEA model were essentially proxies of physical measures: "Since it is very difficult to obtain the

actual data on the number of employees, units of raw material, etc. financial data is used as the next best alternative in this study" (Saranga, 2009, p. 709). However, most Financial DEA literature remains silent on this issue. Hence, the typology classifies models according to the observable accounting variables and the corresponding production elements that indicate the construct being measured.

Productive efficiency

Productive efficiency follows the definition of technical efficiency defined by Farrell (1957), as discussed in section 2.2. Financial DEA models measuring productive efficiency have a distinguishable feature that the models closely relate to a specific underlying production process. In these Financial DEA models, the accounting variables proxy the production elements closely. The typology identified ten models measuring productive efficiency, and they were all in the banking sector. Due to the nature of the banking industry, a significant proportion of variables are monetary (Benston, 2004). Therefore, the productive elements are in the unit of dollars by nature when measuring the technical efficiency of the physical production process.

Financial DEA models measuring productive efficiency treat financial institutions as providers of services, such as transactions and processing of documents in the banking sector. Under this view, the inputs are physical measures, such as labour and capital. The outputs are the number and types of transactions or documents processed. However, since the physical measures are generally proprietary and not available, accounting measures are used as proxies. For example, labour is proxied by personnel costs, e.g. Leightner and Lovell (1998), materials are proxied by funding, e.g. Al-Sharkas et al. (2008) and operating expenses, e.g. Giokas (2008) in banking. On the output side, the transaction and the number of services are proxied by loans and deposits, e.g. Noulas et al. (2008).

Intermediation efficiency

In intermediation efficiency studies, financial institutions are treated as intermediating funds between savers and investors. The typology identified 34 models of this category. To measure the efficiency, inputs are the funds and the interest costs, and outputs are loans and investments, e.g. Athanassopoulos (1997). Due to the special nature of banks, the accounting variables here are direct measures of the intermediation production process.

Historically, intermediation efficiency and profitability efficiency models (discussed in section 4.2.2.2) have been used indiscriminately in the banking sector, e.g. Athanassopoulos (1997). However, with the development of Financial DEA research, studies examining profitability efficiency have separated from intermediation efficiency studies. However, the two DEA models appear similar on occasion, especially when influenced by data availability (Paradi et al., 2011). For instance, Athanassopoulos (1997) incorporated mostly expenses and revenues into the Financial DEA model. However, since the model also covered the number of loans and deposits, this study is categorised into the intermediation efficiency category. By comparison, Giokas (2008) aimed to measure intermediation efficiency but only incorporated expenses to generate revenues and therefore this model was categorised as a profitability efficiency model.

Operational efficiency

In operational efficiency studies, Financial DEA models measure aggregated production processes at the firm level. DEA is a black box by nature in that it does not need to know the process by which DMUs become the most efficient (Rouse et al., 2010). Operational efficiency is different from the productive efficiency. The distinction is that the productive efficiency can only measure one specific production process whereas operational efficiency models can measure the productivity of multiple processes. The typologies identified 54 studies measuring operational efficiency.

For example, Financial DEA models measured dairy-manufacturing firms' operational efficiency without further distinguishing the differences between production lines, e.g. Aparicio and Kapelko (2018). Instead, the Financial DEA models used the cost of labour, the cost of material, and fixed assets as inputs and the revenues as the output. However, dairy-manufacturing firms cover various production activities such as processing raw milk and producing a range of milk products, such as liquid milk, frozen milk products, and fermented milk products (Minj et al., 2020).

In the Financial DEA literature, the indicator of labour is commonly operationalised by remuneration, e.g. Worthington and Hurley (2002). The indicator of materials is commonly operationalised by the cost of materials in the manufacturing sector, e.g. Piesse and Thirtle (2000). The indicator of capital is commonly operationalised by property plant and equipment, e.g. Shiu (2002). Finally, the volume of finished goods is commonly

operationalised to the cost of goods sold with an inventory adjustment to capture the physical volume on the output side. This treatment can avoid the influence of price margin to closely proxy the physical output (Harrison & Rouse, 2016). In the Financial DEA literature, this approach is not commonly observed. Instead, sales with inventory adjustments are used as accounting proxies, e.g. Kapelko et al. (2016).

Financial DEA models categorised to the operational efficiency category do not include the full set of indicators discussed above (i.e. labour, material, and capital). Researchers have omitted the measurement of material in industries that do not have a manufacturing process. For example, neither Zhou et al. (2008) nor Min and Joo (2006) incorporated a specific variable to proxy materials when measuring the operational efficiency of logistic providers due to the nature of the industry. Similarly, Banker et al. (2002) did not incorporate a variable for capital when measuring the operational efficiency of accounting firms.

In some cases, aggregated accounting variables of the production elements are used. For example, Joo et al. (2010) measured the productivity of the retail industry, which is a labour-intensive sector. Even though the study omitted a specific accounting variable for labour, the authors incorporated the accounting variable for selling, general and administrative expenses (*SGA*) instead. Similarly, Mhatre et al. (2014) and Saranga and Phani (2009) omitted the accounting variable for capital when measuring department stores and the pharmaceutical industry. However, they incorporated aggregated variables (*SGA* and cost of production and selling, respectively) as a remedy.

However, it is worth noting a number of studies incorporated detailed accounting variables to proxy all three production elements: labour, material, and capital, e.g. Aparicio and Kapelko (2018). These studies either benefit from a comprehensive database, where more detailed information is required under the accounting regulations, or the researchers had access to unique or hand-collected financial data.

4.2.2.2. <u>Benchmarking</u>

When Financial DEA is used as a benchmarking tool, it forms a best-practice frontier. In contrast to the productive efficiency category (section 4.2.2.1), the production process does not closely underpin the variable selection. In this type of Financial DEA model, accounting variables are not designed to proxy factors of production. Instead, the models maximise

selected outputs, which are the "more-the-better" attributes; and minimise the inputs, which are the "less-the-better" attributes (Cook et al., 2014).

Financial efficiency

The construct of financial efficiency was defined as the efficiency of minimising financial resources while maximising the financial outputs (Demerjian et al., 2012; Demerjian et al., 2013). The majority of Financial DEA literature from accounting journals appears to measure the construct of financial performance. In the typology, 102 models were identified in this category (Appendix 1).

A seminal work by Demerjian et al. (2012) incorporated a range of financial resources as inputs and sales revenue as the output. Unlike the Financial DEA models used to measure productive efficiency (section 4.2.2.1), where inputs and outputs are relatively closely related to productive elements in an underlying production process, models measuring financial efficiency only relate to production processes remotely. Generally, resources (input side) include indicators such as assets and expenses, e.g. Baghdadi et al. (2018). On the output side, indicators to maximise are commonly sales revenue, and occasionally researchers include profit as a measure of revenue quality (Harrison & Rouse, 2016).

Another group of research benchmarks financial performance based on the DuPont ratio (Soliman, 2008). The DuPont ratio uses the return of equity (ROE) to decompose the overall financial performance into three dimensions: (a) the asset utilisation efficiency, (b) the profitability efficiency, and (c) the capital structure efficiency, e.g. Feroz et al. (2003). In the Financial DEA models, the inputs and outputs are similar to Demerjian et al. (2012), with equity as an additional input. Researchers also use Financial DEA models to benchmark selective subcategories, such as profitability efficiency, funding efficiency, and asset efficiency, which will be introduced in detail later in this section.

Profitability efficiency

The construct of profitability efficiency is one dimension of the overall financial efficiency construct. It narrows the financial efficiency to include expenses as inputs and revenues as outputs only. In the typology, 21 models were identified in this category. This type of Financial DEA model is often found in the banking sector, where the inputs are interest expenses and the outputs are interest revenues, e.g. Kaffash et al. (2018).

Marketing efficiency

Marketing efficiency is defined as "the extent to which a firm can minimise advertising inputs…while maximising the outputs…which convert ultimately into sales revenue" (Rahman et al., 2019, p. 621). Generally, these models used advertising expenses as inputs and sales as outputs, e.g. Akdeniz et al. (2010). However, some models also incorporated firm size, e.g. Rahman et al. (2016), and brand value, e.g. Rahman et al. (2018) as outputs to maximise since marketing activity can arguably lead to firm size expansion and increased brand value. Models of this category appear six times in the Financial DEA literature typology.

Funding efficiency

The construct of funding efficiency is another dimension of the overall financial efficiency discussed above. It is a measure of the efficiency brought to a firm through financial leverage. These models include different mixes of capital sources to determine their influence on revenue. In the Financial DEA literature, only four models measuring capital efficiency were identified, e.g. Oberholzer et al. (2017). All models included equity as an input. However, liability was also incorporated by two models, e.g. Smith (1990), but only Smith (1990) incorporated expenses relating to funding sources (interest expenses and tax expenses). On the output side, the models used different indicators. Oberholzer (2014) used assets as outputs; Oberholzer et al. (2017) used sales and profits as outputs; Smith (1990) used net income as the only output.

Asset efficiency

The construct of asset efficiency is measured by including assets as inputs and revenue as outputs. It is defined as the incremental efficiency benefiting firms from owning various types of assets. In the Financial DEA literature, researchers further classify the asset efficiencies to total assets efficiency, e.g. Min and Joo (2009) and current asset efficiency, e.g. Joo et al. (2011), according to a timeline. Six studies that measured asset efficiency were identified: four measured the total asset efficiency, and two measured the current asset efficiency.

Compensation efficiency

The compensation efficiency models measured the efficiency of compensating employees. Five studies were identified in the literature for this type of efficiency, and all the studies focused on executives' compensation. One group of studies took the companies' perspective and used the models to minimise the compensation to maximise the firm resources, such as assets, equity, and revenues, e.g. Oberholzer (2014a). However, the other group of studies took the perspective of executives and used the models to minimise the firms' resources and to maximise the executives' compensation, e.g. Bowlin and Renner (2008). The mixed model construction indicates that the compensation efficiency does not relate closely to a production process. Instead, the models are designed for benchmarking purposes, and the attributes can be maximised or minimised depending on the researcher's objective.

Share market efficiency

The models discussed above focus on internal measures of firm performance. By contrast, share market efficiency is an external measure of financial performance. The Financial DEA models in this category minimise revenues and maximise market values. Some studies used the earnings per share to measure the market value, e.g. Ho et al. (2009). Some Financial DEA models use assets, liabilities, equity, and expenses as the internal resources to minimise, e.g. Frijns et al. (2012). Alternatively, other models used free cash flow as the internal resources to minimise, e.g. Kuo et al. (2020). This category of models incorporates the market value of equity or share performance as one of the outputs, e.g. Frijns et al. (2012). The literature either treats the share market performance as an integral part of overall financial performance Frijns et al. (2012) or as a second-stage ultimate goal following the profit maximisation, e.g. Seiford and Zhu (1999).

In summary, the Financial DEA models identified in the typology were categorised into 12 categories based on the features of the 248 Financial DEA models.[22] Firm performance measured by Financial DEA can be broadly viewed as relating to either productive efficiency, based on some underlying production process, or benchmarking of financial efficiency. When measuring productive efficiency, the accounting variables in Financial DEA models closely relate to production elements in the underlying production process. On the other hand, when

[22] Two models were identified that measured the efficiency of governments in reaching performance goals. However, considering the small quantity of the models discovered, the typology categorised the two models into "others" category.

benchmarking firm performance, the accounting variables only relate to underlying production processes at a distance. Rather, the inputs are the attributes to minimise, and the outputs are the attributes to maximise firm performance. The next section will discuss the classification of indicators and variables based on the 12 dimensional constructs.

4.2.3. **Categories of Indicators and Accounting Variables**

As noted in section 2.5.1, indicators are the signs of the presence or absence of constructs. In Financial DEA, the dimensional constructs are represented by indicators, which are further operationalised into accounting variables. This section extends the discussion in section 4.2.2 by discussing how indicators and accounting variables are utilised to measure different dimensional constructs.

The review of the literature identified nine indicators used in Financial DEA models. Among the nine indicators, four relate to the productive efficiency models (section 4.2.2.1), including labour, material, capital, and production volume. These four indicators will be discussed in section 4.2.3.1, and these factors link to the factors of production according to the view of economics. The other five indicators relate to the financial efficiency benchmarking model (section 4.2.2.2), including assets, liability, equity, expenses, and revenue. These five indicators will be discussed in section 4.2.3.2, and they are the main class of elements from the accounting equation: *Assets + Expenses = Liabilities + Equity + Revenue*.

4.2.3.1. Indicators for productive efficiency

Labour

The indicator of labour represents the human resource in the production process. Together with the indicators of material and capital, the indicator of labour is used by the Financial DEA models to measure the construct of productive efficiency. The labour indicator is conventionally measured physically as full-time equivalent (FTE) employees in conventional DEA to represent an employee's workload.

Table 4 - 3 Accounting Variables for Indicator of Labour

Variables	Frequency
labour cost	21
staff cost	11
remuneration	9
wages and salaries	9
personnel cost	8
compensation	3
employee expenses	3
others	10

Table 4 - 3 summarises the accounting variables that appear in the Financial DEA models to operationalise the indicator of labour. There is a wide variety of variables used to operationalise labour. For example, some researchers included only labour directly involved in the production process, e.g. Yu et al. (2014). In contrast, some researchers used salaries and wages to measure the overall labour expense, including all employees, e.g. Joshi and Singh (2010). However, not all employees in the accounting measure contribute to the production process directly, such as administrative and managerial employees. One measurement issue that could arise during the operationalisation is that the accounting measures do not precisely measure the indicator of labour directly involved in a production process. As a result, the accounting measure is impacted by random measurement errors compared with physical measures (i.e. number of direct labour). This research design may reduce the quality of Financial DEA results reflecting the underlying production process.

Material

The indicator of material is included in the Financial DEA literature classified as the measurement of productive efficiency and represents the resource from which products are produced.

Table 4 - 4 Accounting Variables for the Indicator of Material

Variables	Frequency
cost of goods sold	68
cost of materials	30
cost of inventory	5
cost of energy	5
others	4

As summarised in Table 4 - 4, the indicator of raw material is operationalised by various accounting variables. The most frequently used accounting variable is the cost of goods sold (*COGS*). The *COGS* can proxy material in merchandising firms, e.g. Joo et al. (2011). However, *COGS* is not considered a good proxy for manufacturing due to different business processes, where *COGS* is often used to measure output volume (Harrison & Rouse, 2016). For the manufacturing sector, material is ideally measured by the cost of materials, e.g. Piesse and Thirtle (2000). The cost of energy is included in industries that feature heavy energy consumption, such as the electric utility industry, e.g. Delmas et al. (2007) and the pharmaceutical industry, e.g. Mazumdar (2013).

Capital

The capital indicator refers to the manmade products used in production processes to generate the source of income (Hetico & Marcinko, 2007). In the Financial DEA literature categorised as measuring productive efficiency, the indicator of capital is operationalised using various accounting variables. As illustrated by Table 4 - 5, the most frequently used variable is the net property plant and equipment (NPPE), e.g. Cheung et al. (2017); followed by the depreciation expense, e.g. Zelenyuk and Zheka (2006); followed by the gross property plant and equipment (GPPE), e.g. Yu et al., (2014). However, 41 models only stated using the property plant and equipment (PPE) in the general form without specifying variable used, e.g. Al-Sharkas et al. (2008).

Table 4 - 5 Accounting Variables for the Indicator of Capital

Variables	Frequency
net property plant and equipment	60
property plant and equipment	41
depreciation expense	13
gross property plant and equipment	9
capital expenditure	8
others	18

The variables identified in the literature illustrated five different measurement choices that needed to be considered in relation to capital. First, the choice between flow form and stock form accounting variables incorporates different time frames. Flow form variables refer to accounting variables taken from an income statement, which capture income and expenses

incurred over a specified period. The accounting variable under the flow form is depreciation, defined by the International Accounting Standard (IAS) 16, paragraph 6 as "the systematic allocation of the depreciable amount of an asset over its useful life." (IASB, 2020). By comparison, the stock form variables (i.e. property plant and equipment, gross property plant and equipment, net property plant and equipment, and depreciable value) refer to accounting variables taken from a balance sheet, showing a snapshot of the financial position at a particular point in time. Therefore, it is unlikely that the two forms of accounting variables will provide equivalent measures of capital. For example, if a firm has a high stock of capital but only uses 10% of its capacity, its depreciation will not be proportionally equivalent to the stock value.

Moreover, inconsistency in capital measurement can arise between flow form and stock form due to non-depreciable property, such as land, because it is considered to have an infinite useful life. As a result, no agreed approach in the literature supports either the flow form, e.g. Joo et al. (2010), or the stock form, e.g. Kapelko et al. (2014), as better capital measures. However, most identified Financial DEA studies use stock form variables: 69 models used stock form and only 21 models used flow form variables within the models that specify accounting variables.

Second, accounting variables can use different value bases. In particular, there are two accounting value bases used to measure capital with stock form: gross value base (gross property plant and equipment) or net value base (net property plant and equipment). The gross value base, gross PPE (or cost of PPE), is defined by the IAS, paragraph 6 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" (IASB, 2020). The net value base, net PPE (or carrying amount, or book value of PPE), is defined by the IAS paragraph 6 as "the amount that an asset is recognised after deducting any accumulated depreciation and accumulated impairment losses" (IASB, 2020). Most studies identified did not specify which value base was chosen. However, net PPE is the common choice in quantitative financial accounting research from an accounting reporting perspective because it is a mandatory reporting item in the balance sheet. By comparison, gross PPE is only reported in the notes to the balance sheet.

Third, capital in Financial DEA was operationalised to accounting variables on either historical base or fair value base. According to the IASB conceptual framework paragraph

6.5, the historical cost of "an asset when it is acquired or created is the value of the costs incurred in acquiring or creating the asset, comprising the consideration paid to acquire or create the asset plus transaction costs" (IASB, 2018). In contrast, the fair value is defined by the IASB conceptual framework paragraph 6.12 as "the price that would be received to sell an asset, or paid to transfer a liability, in an orderly transaction between market participants at the measurement date" (IASB, 2018). However, in the literature identified, only a few studies used a fair value base to measure capital. For instance, Mazumdar and Rajeev (2009) used the fair value of PPE to proxy the capital used in the production process of pharmaceutical firms in India. Moreover, Rodríguez-Pérez et al. (2011) used historical-based and fair-value-based PPE in the insurance industry in Spain.

Fourth, the accounting variable to measure capital used in the literature includes accrual basis and cash basis variables. According to the IASB conceptual framework paragraph1.17, the accrual basis is defined as where:

> accounting depicts the effects of transactions and other events and circumstances on a reporting entity's economic resources and claims in the periods in which those effects occur, even if the resulting cash receipts and payments occur in a different period. (IASB, 2018)

In accounting practice, both income statement and balance sheet variables are on an accrual basis, including the stock form and flow form discussed above. However, in cash basis accounting, the IASB conceptual framework paragraph 1.20 stated that "information about a reporting entity's cash flows during a period also helps users assess the entity's ability to generate future net cash inflows and assess management's stewardship of the entity's economic resources" (IASB, 2018). Specifically, the cash basis accounting variables are taken from cash flow statements. A few of the Financial DEA studies identified used the cash basis variable, capital expenditure (*CAPEX*), to operationalise the capital indicator in measuring the productivity construct. Only eight models used *CAPEX* to measure capital in the production process as summarised in Table 4 - 5, e.g. Saranga (2009).

Fifth, the ownership differences between owning and leasing introduces another choice of accounting variable to measure capital. Leasing essentially means that a firm can use capital without ownership. There are two types of leases: (a) operating lease and (b) finance lease. An operating lease is defined by the International Financial Reporting Standard [IFRS] 16,

paragraph 62, as a lease that "does not transfer substantially all the risks and rewards incidental to ownership of an underlying asset" (IASB, 2016). A finance lease is defined as by the IFRS 16, paragraph 62, as a lease that "transfers substantially all the risks and rewards incidental to ownership of an underlying asset" (IASB, 2016).

Historically, the accounting treatment of these two types of leases is different. The operating lease allows firms to have the right to use capital without recording it in the balance sheet. Instead, the usage amount is recorded as an operating lease payment. By comparison, the finance lease requires firms to record the lease on the balance sheet. However, the difference between the two treatments was resolved when the adoption of IFRS 16 became mandatory in 2019. After adopting IFRS 16, both types of leases are recorded on the balance sheet, which is essentially the same as GPPE. However, the inconsistency between the operating lease and finance lease (GPPE) can still cause potential issues for panel data. For example, in Financial DEA, most of the studies did not discuss the difference in ownership types and ignored the operating lease capital when there were different treatments between leasing and owning. Only Demerjian et al. (2012) and Demerjian et al., (2013) and the studies citing their model acknowledged this capital type, e.g. Bonsall IV et al. (2017).

Production volume

In the Financial DEA literature, 55 models were identified using revenue to measure the output volume of an underlying production process, 47 models were from the general sector, and eight models were from the financial sector.[23]

In the general sector, revenue can differ from the output volume due to inventory changes and pricing margins (Harrison & Rouse, 2016). Without adjustment towards output volume, sales can lead to random measurement errors to productive efficiency, e.g. Aparicio and Kapelko (2018). In the Financial DEA literature identified, six models used revenue adjusted with inventory changes to measure the output volume, e.g. Kapelko et al. (2016). In the financial sector, eight models used revenue to measure the income of providing services, e.g. Giokas

[23] Revenue and sales can be used interchangeably for most of the cases. In a very rare case, revenue could be higher than sales since revenue may include supplementary income sources other than the main business activity of selling products and services. In this study, revenue, sales, gross income, and operating income are treated the same.

(2008). However, in financial sectors, models (10 models) also used the value of loan issued to measure the output volume, e.g. Wu et al. (2016).

4.2.3.2. Indicators for benchmarking

Assets

Assets are used in the financial efficiency DEA models as an indicator for the construct of financial performance. An asset is defined by the IASB conceptual framework, paragraph 4.3, as "a present economic resource controlled by the entity as a result of past events" (IASB, 2018). In the Financial DEA literature, assets are treated as a resource to generate revenue when benchmarking financial performance (Demerjian et al., 2012; Demerjian et al., 2013). Within the construct of financial performance, different dimensional constructs are used by the Financial DEA models. Various accounting variables are used in the literature to operationalise assets. The main difference can be found in (a) short-term versus long-term performance and (b) tangible versus intangible assets.

First, short-term versus long-term financial performance can be captured by current assets and non-current assets, respectively. The main decision relies on whether, in a financial year, the resource is transformed into economic benefit. The variable of current assets is expected to be consumed in a year and usually includes variables such as inventories, account receivables, cash, and cash equivalents. Some Financial DEA studies cover all three subcategories of assets: fixed assets, current assets, and intangible assets, e.g. Edirisinghe and Zhang (2008). Some studies only cover selective subcategories of assets, e.g. Demerjian et al. (2013).

Second, long-term financial performance is captured by non-current accounting variables. The variable of non-current assets is the resources to be consumed in the future financial years—the common subcategories link to the classification of tangible and intangible assets. The tangible assets are also known as fixed assets or PPE. However, in contrast to the fixed assets discussed in section 4.2.3.1, the measurement errors here may alter the domain of the construct benchmarked. For example, replacing current assets with fixed assets may change the financial performance from short-term to long-term.

Intangible assets are defined by the IAS 38, paragraph 8, as "an identifiable non-monetary asset without physical substance" (IASB, 2020). In the Financial DEA, only 71 studies

included intangible assets, such as goodwill and research and development expenses, e.g. Bonsall IV et al. (2017). To benchmark financial efficiency, the indicator selection represents industry features. For example, Wang et al. (2020) incorporated research and development expenses (R&D expense) and the other intangibles to measure the financial performance of the telecommunication industry.

Liabilities

Liabilities are defined as "a present obligation of the entity to transfer an economic resource as a result of past events" according to the IASB conceptual framework, paragraph 4.26 (IASB, 2018). In the Financial DEA literature, 35 models incorporate liabilities as inputs, 24 are from the financial sector, and 11 belong to the general sector, reflecting the importance of liabilities in the banking production process where liabilities are used to fund loans and generate income.

In the general sector, the construct of liabilities follows the accounting definition, representing economic resources that could be utilised but are not owned by firms. Among the 11 models, which measure the financial performance in general industries, seven models used liabilities and equity to replace assets as the resource measurement, e.g. Wang et al. (2017). The other four models included all three (assets, liability, and equity) as inputs. That is, they included both the resources used (assets) and the funding sources of those assets (liabilities and equity), e.g. Frijns et al. (2012). Also, the operationalised accounting variables are classified according to the time frame. As noted in the assets section above, the selection depends on individual research aims, whether a short-term or a long-term view of financial performance is the research goal.

Liabilities more frequently appear in the financial sector when measuring productive efficiency. There are various forms of variables to measure liabilities on the input side. Most commonly, liabilities are operationalised as deposits, which are savings from lenders, representing the fund resource in calculating intermediation efficiency, e.g. Hsiao et al. (2010). Or provisions for loan losses are included as negative outputs, e.g. Hadad et al. (2011).

Equity

Equity is defined as "the residual interest in the entity's assets after deducting all its liabilities" in the IASB conceptual framework paragraph 4.63 (IASB, 2020). In the Financial DEA literature, the indicator of equity appears in both the financial sector (25 models) and general sectors (31 models). In addition, it shows up as both inputs (31 models) and outputs (25 models) in the reviewed Financial DEA models.

In general sectors, equity appears on the input side in the Financial DEA models to measure financial performance, in accordance with the DuPont model. As noted in section 4.2.2.2, the DuPont model decomposes the ratio of return of equity (ROE) into three dimensions, including the profitability efficiency, asset utilisation efficiency, and capital structure efficiency, in the order as illustrated below (Soliman, 2008):

$$ROE = \frac{Net\ Income}{Sales} \times \frac{Sales}{Assets} \times \frac{Assets}{Equity}$$

In the Financial DEA literature, the assets, equity, and expenses are commonly used as the inputs to produce sales when measuring firm performance and, specifically, the efficiency of generating sales from resources, e.g. Feroz et al. (2001).

There is a range of equity variables that can be used as an input. Stockholder equity is the most common choice (four models). On the output side, the market value of equity is used in the Financial DEA literature to measure the share market performance as a dimension of financial performance. Eight models that had equity as both an input and output in the same model were identified. Seven models[24] used the book value of equity as an input to generate the market value of equity, e.g. Kweh et al. (2014). Similarly, researchers used two-stage models to measure the share market performance construct after measuring the construct of profitability efficiency in the first stage, e.g. Seiford and Zhu (1999).

In the financial sector, equity also appears as both input (15 models) and output (10 models). As an input, equity measures the monetary resources used in productive efficiency, e.g. Al-Sharkas et al. (2008) or intermediation efficiency, e.g. Chen et al. (2014). As an output,

[24] Smith (1990) used the book value of equity (input) to generate retained earnings (output) (Smith, 1990), to benchmark the overall financial performance.

equity exists in the insurance industry only, representing the financial benefit received in reserves, e.g., Kweh et al. (2014).

Expense

In the Financial DEA literature, to benchmark firm performance, the indicator for expense appears in both the financial (22 models) and general sectors (103 models). In the general sector, expenses are used as attributes that are to be minimised for financial performance benchmarking. For instance, the Financial DEA based on DuPont ratios incorporated the operating expense variable as inputs, e.g. Ho et al., (2009) to measure general financial performance. In the financial sector, expenses are used to measure profitability efficiency, where the interest expenses are to be minimised while the interest income is maximised, e.g. Kaffash et al. (2018).

Revenue

In the Financial DEA literature, the indicator of revenue appears 142 times in total, 24 times in the financial sector and 118 times in the general sector. It exists mainly on the input side (131 models) and occasionally on the output side (11 models).

In the general sector within the literature measuring the construct of financial firm performance, two ways of using revenue were identified: (a) DuPont ratio, e.g. Feroz et al. (2001) and (b) output maximising dimension, e.g. Demerjian et al. (2013).

In the financial sector, the indicator of revenue measures the profitability efficiency (20 models), which minimises the interest expenses and maximises the interest revenue, e.g. Kao and Hwang (2008). The other four models used revenue as the attribute to maximise and measure financial efficiency, e.g. Wanke et al. (2020).

When revenue is on the input side, the Financial DEA models measured constructs such as (a) executives' compensation efficiency by taking the executive perspective, e.g. Bowlin and Renner (2008); (b) Share market efficiency, e.g. Kuo et al. (2020); or (3) Financial efficiency by generating profits, e.g. Tsolas (2011).

In summary, this section (section 4.2.3) identified nine indicators for the dimensional constructs found in section 4.2.2. There are four indicators used for productive efficiency following a reflective model: labour, material, capital, and output volume. There are five

indicators used for benchmarking, following a formative model: assets, liabilities, equity, expense, and revenue.

4.2.4. The Breadth of Firm Performance in Financial DEA

The sections above summarised nine indicators and 12 dimensional constructs identified in the Financial DEA literature. At the operational level, various combinations of the nine indicators illustrate the 12 dimensional constructs following either reflective or formative modelling (section 4.2.2 and section 4.2.3). At the conceptual level, the relationship between dimensional constructs and firm performance phenomenon also follows either reflective or formative modelling (section 2.5.2.3), relating to two views of the performance construct.

Using the lens of reflective modelling, firm performance is a general phenomenon with implicit meaning. The Financial DEA literature in this category treats the dimensional firm performance constructs (section 4.2.2) and the firm performance phenomenon interchangeably (section 2.4.1). For example, some researchers used the dimensional construct "productive efficiency" and treated this construct interchangeably with firm performance, e.g. Charnes et al. (1990). Another group of researchers used the dimensional construct "financial efficiency" and treated it interchangeably with firm performance, e.g. Baghdadi et al. (2018).

Using the lens of formative modelling, firm performance is a combination of separate dimensional constructs. The Financial DEA literature in this category conceptually defined separate dimensional constructs (section 4.2.2) and combined these dimensional constructs to define the firm performance phenomenon (section 2.4.2 and section 2.4.3).

For example, in the financial sector, researchers using this approach identified multiple separate Financial DEA models to measure different dimensional constructs, including intermediation efficiency, profitability efficiency and productive efficiency, e.g. Kaffash et al. (2018). Researchers have measured the total asset efficiency and current asset efficiency in general sectors by incorporating accounting variables of different time frames, e.g. Joo et al. (2011).

To sum up, this section discussed the measurement modelling relationships between the identified dimensional constructs and the firm performance phenomenon at the conceptual level. The relationships between dimensional construct and the firm performance

phenomenon are of two categories: reflective modelling, where firm performance is treated as a common latent phenomenon and formative modelling, where firm performance is treated as a composite latent construct.

4.3. Methodological Issues in Financial DEA Application

This section synthesis the methodological issues in DEA application (section 2.2.4 and measurement models (section 2.5) with the Financial DEA literature typology in section 4.2, to discuss the methodological that potentially arise in Financial DEA application. These methodological issues include homogeneity (section 4.3.1), discriminatory power (section 4.3.2), non-negative variables (section 4.3.3), and selection of variables (section 4.3.4) and errors in measurement errors (section 4.3.5).

The issues discussed in this section are based on the analysis of the Financial DEA literature. Financial DEA is located at the nexus of DEA, accounting information, and performance measurement. The feature of accounting information (section 2.3), which is different from the conventional DEA's physical measures, can emphasise methodological issues relating to Financial DEA. Also, the measurement models (section 2.5) emphasise methodological issues arising during the Financial DEA research process.

4.3.1. **Homogeneity**

DEA has a series of assumptions related to the homogeneity of the DMUs under assessment. The homogeneity assumption requires that DMUs are similar as follows. First, DMUs are assumed to conduct similar activities and produce comparable goods or services from similar resources. Second, the operating environments of DMUs need to be similar since the external environment could impact the overall performance of organisations (Dyson et al., 2001). Therefore, financial DEA models need to reflect similar activities or processes and operate in similar environments such as industry, location, and period. This is so heterogeneous factors do not bias results (Joo et al., 2011).

In Financial DEA, some researchers have argued that all accounting variables can be combined since they are in monetary form (Dyson et al., 2001). However, according to Farrell (1957), if the price factors are not stable, the technical efficiency estimated by overall efficiency will vary according to the price variation. This is because price factors are

influenced by pricing strategies within firms and by the demand and supply status of the free market.

In the Financial DEA typology (section 4.2), researchers have applied Financial DEA models across various DMUs creating potential homogeneity issues. For example, some researchers have calculated firm efficiencies using a single Financial DEA model across multiple industries using the Fama-French classifications (Baghdadi et al., 2018; Demerjian et al., 2012); firms across multiple industries listed on Standard and Poors (S&P) (Bowlin & Renner, 2008), the New York Stock Exchange (NYSE), Nasdaq Stock Market (Frijns et al., 2012), the Ukrainian stock exchange (Zelenyuk & Zheka, 2006) and China Securities Regulatory Commission (Wang et al., 2017). The heterogeneity among DMUs due to the variations in industry, operating activities, goods and services, operating environment and prices may not be aligned.

4.3.2. **Discriminatory Power**

Discriminatory power describes the ability of DEA to distinguish efficient DMUs from inefficient ones. In DEA, there is a relationship between sample size (number of DMUs) and model size (number of inputs and outputs), which can influence the discriminatory power of DEA models, as discussed in section 2.2.4. Discriminatory power is low when efficiency scores are spread in a narrow range around full efficiency. Low discriminatory power is usually due to the inclusion of an excessive number of variables in a DEA model (Cook et al., 2014).[25]

In Financial DEA, the risk of low discriminatory power is highlighted where researchers include relatively large models (number of inputs and outputs). In the typology of Financial DEA literature (section 4.2), relatively large models are used by several researchers. For example. Demerjian et al. (2012) used seven inputs and one output. The large size of inputs and outputs could potentially reduce the discriminatory power. Especially when the sample size (number of DMUs) of an industry is relatively small. For example, in Demerjian et al. (2012, p. 1236), the smoking industry has 268 observations in total over 29 years, which has on average 9.24 DMUs per year.

[25] This could also happen in groups of DMUs that have high structural efficiency. That is homogeneous set with very similar levels of productivity.

However, the inclusion of multiple financial years without adjusting for inflation or considering the potential change of technology could reduce the comparability between DMUs representing different financial years. For instance, Demerjian et al. (2018) included data from 1980 to 2015; Baik et al. (2013) used data from 1976 to 2008. However, these studies have not discussed whether accounting variables have been adjusted for inflation or treatment of potential technology change.

4.3.3. **Non-negative Variables**

The DEA formula requires non-negative variables to generate feasible results. As suggested by the reseachers, there are two different ways to treat negative variables: (a) if the DMUs with negative variables cover a small portion of the data set, these data can be removed from the group e.g. Bowlin (1995); and (b) the negative values can be translated by adding a constant to the variable for each DMU equal to at least the largest negative value in the sample e.g. Bowlin (1999).

In Financial DEA typology (section 4.2), the accounting variables used are potentially subject to a relatively high proportion of negative values because of the accounting equation (section 4.2.3.2). Only two Financial DEA models incorporated net income in the literature typology, potentially because of this issue. Smith (1990) and Wang et al. (2020) incorporated net income into the Financial DEA models but did not explain the treatment of negative values. Another group of studies transformed the Financial DEA models by applying the accounting equation (*Income = Revenue – Expenses*) to avoid negative values in the DEA model. For example, Feroz et al. (2001) initially constructed a Financial DEA model using the DuPont ratio to minimise total assets, sales, and equity and maximise income. However, the sample showed a high percentage of negative income data points. The authors later refined the model to minimise total assets, equity and expenses and maximise sales, considering that income is the difference between revenue and expenses.

4.3.4. **Selection of Variables**

The selection of variables for DEA determines the factors considered in determining the efficiency of DMUs. In Financial DEA, where researchers applied productive efficiency models, input and output factors should closely reflect the production factors of the underlying production process (Golany & Roll, 1989). The factors of production are defined as "the distinct kinds of goods and services used in production" (Shephard, 1970, p. 13). In

this sense, accounting variables may not provide a perfect measurement of a production process since all measures are in monetary form. The Financial DEA function does not need to relate to a production function precisely in financial efficiency benchmarking models. Instead, a best-practice frontier is built. The input and output variables are chosen to build performance metrics. Mathematically, inputs are to be minimised and outputs maximised; therefore, the less of the former and the more of the latter lead to better performance.

In the Financial DEA typology (section 4.2), models classified as productive efficiency tend to select accounting variables closely related to the production process. For example, Aparicio and Kapelk (2018) measured the productive efficiency of dairy firms by using labour cost, material cost, and fixed assets generating revenues. By comparison, the Financial DEA models classified as financial efficiency benchmarking selected accounting variables based on a performance benchmark. For instance, Demerjian et al. (2012) measured the efficiency of generating revenues by using the cost of goods sold (*COGS*), selling, general, and administrative expenses (*SG&A*), net property plant and equipment (*NPPE*), operating lease, net research and development (*Net R&D*), and goodwill and other intangibles as inputs.

4.3.5. **Errors in Measurement Models**

This section discusses potential measurement errors in Financial DEA modelling based on measurement models (i.e. reflective model and formative model) discussed in section 2.5. Based on the feature of these models, measurement errors could arise

Within Financial DEA measurement models (section 2.5), there are two types of errors: (a) random measurement errors in variables in reflective models; and (b) disturbances in composite constructs in formative models (Edwards & Bagozzi, 2000; Jarvis et al., 2003; Law et al., 1998; Law & Wong, 1999).

First, random measurement errors are defined as the difference between measures and true values (Nunnally & Bernstein, 1994). Random measurement errors are due to the invalidity and unreliability of measures, which could be caused by contaminated constructs or random factors (Mackenzie et al., 2005). In the measurement models, measurement errors exist in reflective models.

In the reflective measurement model between accounting variables and dimensional constructs, random measurement errors arise on variables at the operational level. Random

measurement errors are unique to each variable. They are the part of each variable that is not covariant with the common cause, that is, the underlying dimensional constructs (Edwards & Bagozzi, 2000; Jarvis et al., 2003); for example, Chang et al. (2015) used accounting variables in reflective models designed to measure productive efficiency. Productive efficiency is best reflected by physical measures (Färe et al., 2017; Farrell, 1957); one source of measurement error is fluctuating and potentially unknown pricing information, which will cause variation in the relationships between accounting measures and physical measures.

Random measurement errors arise on dimensional constructs in the reflective measurement model between dimensional constructs and the firm performance phenomenon. Measurement errors are unique to each dimensional construct due to the difference between the dimensional construct and the ideal (or true) dimensional construct. The measurement errors are the dimensional constructs that do not co-vary with the common cause, that is, the underlying concept.

For example, Demerjian et al. (2012) designed Financial DEA constructs meant to reflect the firm performance concept as a whole. However, Financial DEA constructs may contain parts that do not co-vary with the intended firm performance construct. Since firm performance is a sophisticated concept, there are differing views within the literature. Therefore, researchers need to clearly define the firm performance concept and be consistent in implementing empirical tests to minimise the potential for random measurement error.

Second, disturbances are defined as part of the construct that facets cannot explain (Jarvis et al., 2003). The disturbance captures the invalidity of the construct, which may be due to the imperfect validity of individual facets or their invalidity as a group due to the failure to include all facets that are decisive aspects of a construct (Mackenzie et al., 2005). Disturbances exist in composite constructs, and the variables here are assumed to be error-free causes of these constructs (Edwards & Bagozzi, 2000). In a formative model, if variables operationalise an incomplete set of indicators, the content validity and construct validity are undermined (Bisbe et al., 2007; Ghauri & Grønhaug, 2010). Since the relevant domain of the construct is not adequately covered by omitted decisive indicators, the meaning of the composite construct changes.

In the formative measurement models between accounting and dimensional constructs, the dimensional constructs are affected by disturbances. In Financial DEA, for example, Joo et

al. (2011) used a range of accounting variables to form dimensional constructs. However, if any indicator was not operationalised to variables, the Financial DEA constructs' domain might have changed. Thus, researchers need a census of the indicators, but not sampling, and researchers need to know to what extent the disturbances lead to errors in Financial DEA studies.

The phenomenon is affected by disturbances in the formative measurement models between dimensional constructs and the firm performance phenomenon. In a formative measurement models, firm performance is a multi-dimensional construct formed by separate dimensional constructs, representing different aspects of firm performance. Ignoring any decisive dimension of firm performance can lead to disturbances in the firm performance constructs. For example, in Paradi et al. (2011), intermediation efficiency and profitability efficiency are argued to capture the overall bank performance, but other performance dimensions may have been omitted. Researchers need to address to what extent the disturbances influence the measurement of firm performance.

4.4. Propositions for Empirical Tests

Based on the Financial DEA typology (section 4.2) and the potential methodological issues in Financial DEA application (section 4.3), this section develops three sets of propositions to test selective issues empirically. The three sets of empirical tests developed in Chapter 5 based on these propositions are the impact of (a) the price factor (section 4.4.1), (b) alternative forms of accounting variable (section 4.4.2), and (c) alternative accounting indicators and variables (section 4.4.3) on the results of Financial DEA.

4.4.1. Proposition One: The Price Factor

As discussed in section 4.3.1, fluctuation of prices could introduce heterogeneity into Financial DEA and influence the results. In business practice, firms usually treat price data as confidential information and do not disclose prices. However, in many cases, Financial DEA studies use aggregated accounting data (e.g. revenues) as proxies to calculate technical efficiency in DEA (Banker et al., 2007; Zelenyuk, 2020). Accounting data can be regarded as the product of physical measures of inputs and input prices adjusted by accounting reporting regulations.

As noted in Chapter 2 (section 2.2.2), originally, DEA developed from productivity theory in economics (Charnes et al., 1978; Farrell, 1957). Financial DEA is distinguished from conventional DEA in that the inputs and outputs are accounting measures rather than physical measures. Farrell (1957, p. 254) defined technical efficiency as "(firms') success in producing as large as possible an output from a given set of inputs". Farrell (1957, p. 259) distinguished allocative efficiency and technical efficiency, stating that "the former measures a firm's success in choosing an optimal set of inputs, the latter its success in producing maximum output from a given set of inputs". Farrell (1957, p. 260) also commented that "price (allocative) efficiency is a measure that is both unstable and dubious of interpretation; its virtue lies in leaving technical efficiency free of these faults". Based on productivity theory, only when all inputs and outputs are physical measures can DEA calculate technical efficiency reliably (Farrell, 1957).

Based on the definition of technical efficiency, the researchers have theoretically expressed concern over using accounting variables to calculate technical efficiency. The researchers stated that aggregated accounting data could not compute technical efficiency since it requires purely physical measures (Portela., 2014). Instead, only overall efficiency can be calculated with aggregated accounting variables if the prices are different across DMUs (Portela, 2014). Only if prices are known to be the same across DMUs can technical efficiency be computed from the overall efficiency and aggregated accounting data (Färe et al., 1990; Färe et al., 2017). In recent literature, no allocative inefficiency is used as a condition to provide unbiased results of DEA with price-based aggregation (Zelenyuk, 2020).

In section 4.2.2, the Financial DEA literature typology found 98 models using accounting variables to proxy productive elements relatively closely (productive efficiency – 10 models, intermediation efficiency – 34 models, and operational efficiency – 54 models). In these models, measurement errors or heterogeneity are introduced by the variation of prices, which relate to the homogeneity issues discussed in section 4.3.1 and the measurement errors discussed in section 4.3.5. In conventional DEA definitions, when all variables are physical measures, the DEA model measures technical efficiency. This proposition aims to empirically investigate whether Financial DEA's results in estimating technical efficiency are impacted when prices vary to lesser or greater extents. Hence the proposition is stated as:

***P_1**. As variation in prices in the accounting data increase, the Financial DEA results provide less convergent validity in estimating physical technical efficiency.*

The results testing this proposition will be reported in section 5.2.

4.4.2. **Proposition Two: Stock and Flow Forms of Accounting Variables**

As discussed in section 4.3.4, the selection of accounting variables determines the factor considered when calculating efficiency. When calculating productive efficiency, the accounting measures may not provide a perfect measurement of a production process since they are in monetary form. Also, the typology of the indicators for productive efficiency identified that for one productive element (i.e. labour, material, capital, and production volume), various forms of accounting variables were used by researchers. For example, in section 4.2, of the 248 Financial DEA models using capital, identified in the literature, 147 contained a measure of capital specified using an accounting variable: most models used *NPPE* (60 models), followed by *DP* (13 models), and only nine models used *GPPE*.[26]

This proposition considers how alternative forms of accounting variables act as proxies for the elements of production in Financial DEA. The proposition focuses on the stock and flow forms of accounting variables used to measure capital. Stock form accounting variables are measured at a specific point in time and represent the value at that time (i.e. accounting variables disclosed in the balance sheet). Flow form accounting variables represent the value over a specific unit of time (e.g. a financial year) and represent an increase or decrease over a period in the sense of rate or speed of use (i.e. accounting variables from the income statement) (Ponta et al., 2018). In economics, the stock-flow consistent (SFC) modelling approach requires consistency in accounting. Every flow implies a change in one or more stocks, and the end-of-period stocks are obtained by accumulating the relevant flows (Nikiforos & Zezza, 2017). This raises the question of whether stock and flow form accounting variables lead to different Financial DEA results when measuring capital in a production process.

Capital can be proxied by either stock form or flow form accounting variables to measure a physical production process. The physical capital can be measured by various property, plant

[26] There are 43 studies which did not specify which from of PPE was used to proxy capital. Eight studies used capital expenditure. And 18 studies used other accounting variables.

and equipment accounting measures disclosed in the accounting statements.[27] According to the IAS16, PPE is defined as "tangible assets that are held for use in the production or supply of goods or services, for rental to other, or administrative purposes and are expected to be used during more than one period" (IAS16, 2019, para.6). The stock from of PPE: (a) gross property, plant and equipment (*GPPE*), is a stock measure of capital, as at a historical point of time; (b) net property, plant and equipment (*NPPE*) is a stock measure of capital, as at a current point of time; and (c) the flow form of PPE is depreciation expense (*DP*), which is the allocation of the depreciable value of capital over a financial year.

GPPE is defined by the accounting standard 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 recognised in accordance with the specific requirements of other IFRSs." (IASB, 2019)[28]

The *GPPE* is a stock form accounting variable as it is the value of capital at a historical point when the capital was purchased. The advantage of *GPPE* as a proxy of capital is that the value of *GPPE* is relatively stable since past transactions verify it. The estimation of depreciation rate, or variations in the operating environment, do not influence the value of *GPPE*. When capital is assumed to be of similar ages across firms, the *GPPE* is not affected by variation in buying powers or inflation. However, the disadvantage of using *GPPE* is that the *GPPE* does not capture the usage over a year or varies significantly from year to year.

Table 4 - 6 Illustrative Example

	Machine A	Machine B	Productivity A	Productivity B
GPPE (cost)	$ 100	$ 100	high	low
DP (usage)	$ 50	$ 10	same	same
NPPE (future value)	$ 50	$ 90	high	low
Output	$ 50	$ 10		

[27] PPE has alternative names, such as tangible assets and fixed assets.
[28] 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” (IAS16, 2019, para.6). However, to restrict the focus of this study, only historical value is considered.

For example, in Table 4 - 6, there are two machines, A and B, with the same cost ($100). Machine A is used more over a year (*DP* = $50) and produces more output (output = $50), and Machine B is used less (*DP* = $10) and produces less output (output = $10). Using *GPPE* as a proxy for the capital will lead to Machine A having higher measured productivity than Machine B. It can be argued this is consistent with an investment view insofar as Machine A and B require the same investment, but Machine A is used more productively.

By comparison, the stock value, as at a current point in time, is captured by the *NPPE* and may be a better proxy for physical capital because it represents the value of the capital after deducting the usage to date. *NPPE* is defined as "the amount at which an asset is recognised after deducting any accumulated depreciation and accumulated impairment losses" (IAS16, 2019, para.6). The *NPPE* is a stock form of accounting variable because it is the capital value, as at the end of the most recent financial year. The advantages of *NPPE* are that it captures both the cost and the accumulated utilisation over past financial years. However, the disadvantage of *NPPE* is that it represents the remaining portion of the asset but not the value used for the most recently finished production period. Further, *NPPE* is a more sensitive variable to accounting choices. Its calculation is influenced by accounting choices related to measuring the stock form (*GPPE*) and the flow form (*DP*).

As shown in Table 4 - 6, Machine A with a higher utilisation rate will lower *NPPE* ($50) due to the higher *DP* ($50). By comparison, Machine B has a higher *NPPE* ($90) due to lower DP ($10). As a result, if *NPPE* is used as the machine's proxy, Machine A has higher measured productivity than Machine B.[29]

However, neither the capacity measured by *GPPE* nor *NPPE* considers the utilisation rate of capacity, which relates to the hours the plant operates as defined by Ray et al. (2020). Instead, the utilisation rate is arguably captured by the *DP*. Depreciation (*DP*) is defined as "the systematic allocation of the depreciable amount of an asset over its useful life" (IAS16, 2019, para.6).[30] The *DP* is a flow form accounting variable since it represents the capital utilisation

[29] Also, for the same piece of machine, assuming the yearly output is constant, the NPPE will reduce in the following year. As a result, the productivity calculated with NPPE for the first year will be lower than in the following year.

[30] According to the definition of depreciation, useful life is defined as “the period over which an asset is expected to be available for use by an entity” (IAS16, 2019, para.6). The “depreciable amount is the cost of an asset or other amount substituted for cost, less its residual value” (IAS16, 2019, para.6). The “residual value of an asset is the estimated amount that an entity would currently obtain from disposal of the asset, after deducting the estimated costs of disposal, if the asset were already of the age and in the condition expected at the end of its

rate in monetary form. The advantage of using *DP* to proxy capital is that *DP* contains information on the utilisation rate of assets. However, the disadvantage of *DP* is that the value is sensitive to the estimation of useful life. In addition, the estimation is subject to managerial estimation and operating environment, which would introduce heterogeneity into the accounting variables and the Financial DEA results.

As shown in Table 4 - 6, the two pieces of machinery have the same cost, but Machine A is utilised more intensively and has a higher *DP* ($50) and shorter useful life ($100/$50 = 2 years, assuming no residual value) than Machine B, which has a relatively lower *DP* ($10) and longer useful life ($100/$10 = 10 years, without residual value). Thus, compared with *GPPE* and *NPPE*, *DP* leads to the measured productivity of Machine A and B, reflecting the differences in the relative usage. Still, it does not reflect the fact that the initial capital investment is the same.

This proposition aims to examine the impact of alternative forms of accounting variables on the Financial DEA results. The stock and flow form accounting variables have been chosen to test this. The choices between stock and flow accounting variables essentially introduce various accounting information and operational factors. For example, to measure the amount of physical capital in a production process, stock form accounting variables used are *GPPE* and *NPPE* and flow form accounting variables used are *DP*. To build propositions, this section analyses the mathematical relationships between *GPPE*, *NPPE* and *DP* and isolates selected accounting and operational factors that impact the measurement of *GPPE*, *DP* and *NPPE*.

First, according to the accounting definition of *DP*, the mathematical equation for its calculation is expressed as below:

$$DP = \frac{Depreciable\ Amount}{UL} \qquad (4-1)$$

Equation (4 - 1) follows the straight-line depreciation method, which assumes the same depreciation amount each year of the plant's useful life (*UL*). Another common depreciation

useful life" (IAS16, 2019, para.6). Depreciation does not normally apply to land since its usefulness and revenue-producing ability generally remains intact over time. There are two main types of depreciation method, being the straight-line method and the diminishing value method. The straight-line method assumes a fixed amount of depreciation every year. The diminishing value method assumes a fixed rate of the opening balance of PPE to be depreciated every year.

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. The test developed in Chapter 5 uses the straight-line depreciation method since it is simple to apply and matches the expenses with revenue when the use of the plant is reasonably uniform throughout the estimated life (Weygandt et al., 2015).

The full expression of equation (4 - 1) is:

$$DP = \frac{GPPE - residual\ value}{UL} \qquad (4-2)$$

In practice, the residual value is a small portion of *GPPE*. For simplicity, this proposition will not consider the heterogeneity of residual values and assumes all assets have a zero-residual value.

Equation (4 - 1) indicates that the depreciable amount is fixed and determined by the past transaction (*GPPE*) for a certain machine. The *UL* and *DP* are related inversely and form an inverse proportional function. As a result, the *DP* value is influenced by the accounting estimation of *UL*. In accounting practice, the *UL* is determined by firms who estimate the length of useful life by comparing it with similar assets (Weygandt et al., 2015).

Figure 4 - 2 illustrates how *DP* is impacted by *UL*. In Figure 4 - 2, the x-axis represents *UL*, and the y-axis represents the value of *DP*. For a certain capital (Capital A), the cost is verified by the past market transaction, and it is fixed ($GPPE_{A.}$). The fixed value of $GPPE_A$ is the area of the rectangle, which is the product of UL_A and DP_A. For the capital, if the estimated useful life increases ($UL_{A'} > UL_A$), the utilisation rate of the capital reduces, which leads to a lower depreciation expense ($DP_{A'} < DP_A$) for each year.

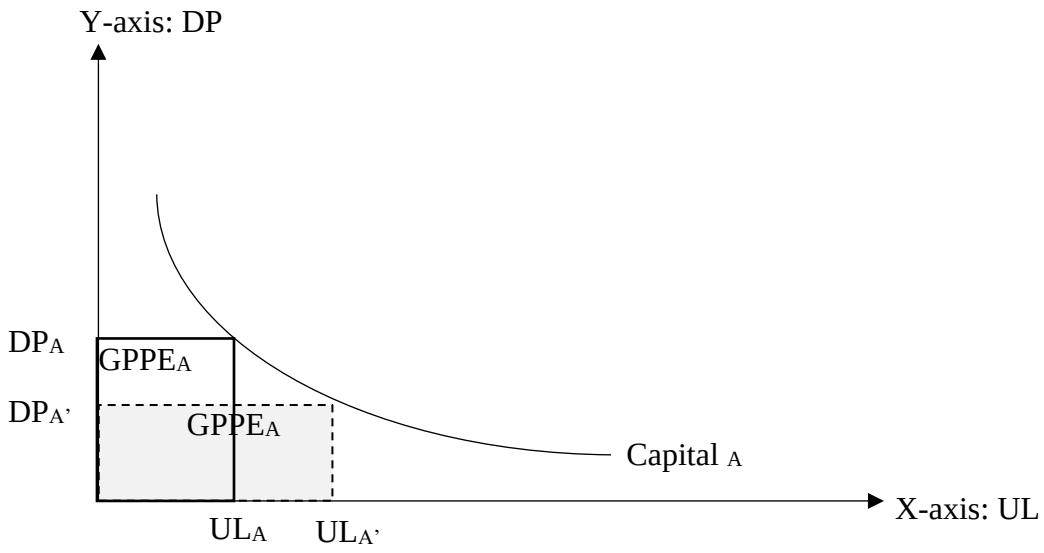


Figure 4 - 1 Mathematical illustration of the relationship between DP and UL

As illustrated by equation (4 - 3), *DP* is essentially *GPPE* times a multiplier, which is the inverse of *UL*. Hence the proposition is stated as:

> ***P₂ₐ.*** *When the cost of the plant is fixed (GPPE), a longer estimated useful life (UL) generates smaller depreciation expenses (DP) and higher Financial DEA efficiency scores when using depreciation expenses (DP) as the measurement of capital.*

NPPE is expressed by the mathematical equation below according to the accounting definition:

$$NPPE = GPPE - AGE \times DP \qquad (4-3)$$

As noted above, the test developed in Chapter 5 uses the straight-line method when calculating the amount of depreciation. Equation (4 - 3) is a mathematical expression of the straight-line depreciation method. According to equation (4 - 3), *NPPE* is influenced by the age of capital (*AGE*) and *DP*. Moreover, *DP* is affected by *UL*, as shown in equation (4 - 2). To isolate the factors that impact the value of NPPE, equation (4 - 2) and (4 - 3) are combined. Equation (4 - 4) isolates the relationship between *NPPE* and the two factors, *UL*, and *AGE*. The rate of *AGE* and *UL* is defined as the relative age rate ($\frac{AGE}{UL}$) for the next proposition.

$$NPPE = GPPE \times \left(1 - \frac{AGE}{UL}\right) \qquad (4-4)$$

$$NPPE = -GPPE \times \frac{AGE}{UL} + GPPE \qquad (4-5)$$

To further understand the relationship between *GPPE* and *NPPE*, equation (4 - 4) is rearranged in the linear form of $y = -ax + b$, the format of a straight line. As illustrated in equation (4 - 5), GPPE is fixed, and NPPE is a function of the relative age rate ($\frac{AGE}{UL}$).

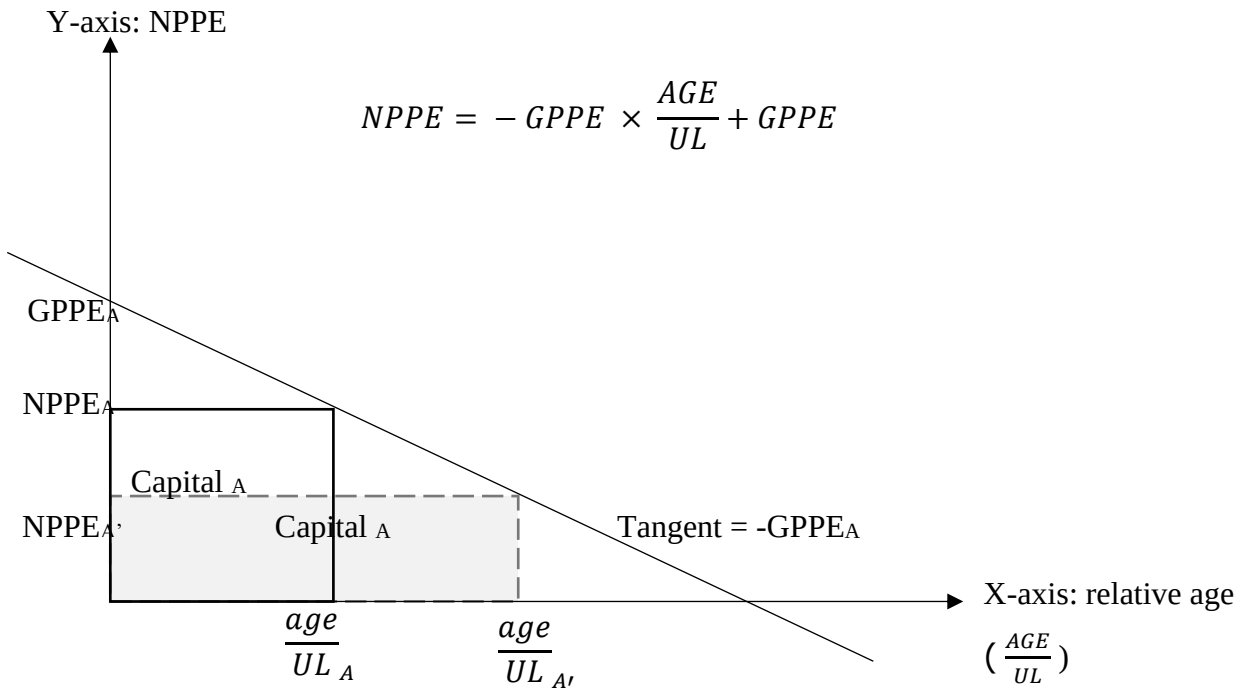


Figure 4 - 2 Mathematical illustration of the relationship between NPPE and Relative Age

Graphically, Figure 4 - 3 illustrates equation (4 - 5), that is, the relationship between the relative age rate ($\frac{AGE}{UL}$), which is the independent variable (x-axis), and *NPPE*, which is the dependent variable (y-axis). The interception of the straight line and the y-axis is the value of GPPE. That is the cost of Capital$_A$ ($GPPE_A$) according to the past transaction. The figure illustrates that, for a certain piece of capital, the higher the relative age ($\frac{age}{UL_{A'}} > \frac{age}{UL_A}$), the lower the net book value of the machine ($NPPE_{A'} < NPPE_A$).

When *NPPE* is impacted by the variation of one factor of interest (*UL* or *AGE*) separately, if the relative rate ($\frac{AGE}{UL}$) is the same; the Financial DEA results will be identical. For simplicity, the proposition treats as alternatives the impact of *UL* and *AGE* on *NPPE*, stated as:

> ***P*$_{2b}$.** *When the cost of a plant is fixed (GPPE), and the relative age rate ($\frac{AGE}{UL}$) is the same; either useful life (UL) or age (AGE) can vary independently and calculate the other variable accordingly. As a result, if using NPPE as the proxy for capital, the Financial DEA results will be identical.*

However, equation (4 - 5) suggests that *NPPE* is a function of relative age rate (AGE/UL) but not individual AGE or UL. Hence the proposition is restated as:

> ***P*$_{2c}$.** *When the cost of a plant is fixed (GPPE), the higher the relative age rate ($\frac{AGE}{UL}$), the lower the NPPE and the higher the Financial DEA efficiency scores if using NPPE as a proxy for capital.*

To sum up, *GPPE*, *DP*, and *NPPE* are alternative accounting variables used to proxy the factor of capital in a production process. Various sources of heterogeneity impact each variable due to accounting choices (e.g. estimation of useful life) and operational characteristics (e.g. age of capital). From a reflective modelling view, these variations may lead to heterogeneous accounting measurements and divergence in the Financial DEA results compared to the underlying physical productivity. However, from a formative modelling view, these variations may add information to the construct needed for managerial decision making. Information such as operational characteristics and accounting choices may provide additional attributes to the construct. The divergence in the Financial DEA construct from the physical productivity will create a hybrid construct of managerial information used for decision-making and productivity. The results of testing these propositions will be reported in section 5.3.

4.4.3. **Proposition Three: Alternative Accounting Variables**

As discussed in section 4.3.4, when Financial DEA models benchmark financial efficiency, variables are chosen based on performance metrics. According to the feature of formative models (section 2.5.2.2), the change of variables may alter the meaning of the construct formed.

This proposition examines the impact of alternative accounting variables on the construct validity of Financial DEA when used for benchmarking purposes under formative modelling. In contrast to the propositions in Proposition Two (section 4.4.2), which examine the impact on Financial DEA when alternative forms of one accounting variable is used for one indicator, Proposition Three (section 4.4.3) examines how alternative accounting variables influence Financial DEA results, when the indicators are fixed or alter from the perspective of formative modelling.

Financial DEA literature has used various accounting variables to measure firm performance. According to the typology in section 4.2, when Financial DEA is used to benchmark firm performance (150 studies), the most common construct measured is financial efficiency (102 studies). These studies generally measure the financial construct as the efficiency of transforming resources into revenues (Baghdadi et al., 2018; Demerjian et al., 2013; Hasan, 2020).

As a benchmarking tool, Financial DEA models use various accounting variables to operationalise indicators and benchmark firm performance. As noted, the indicators are the elements and sub-elements from the accounting equation:

$$Assets + Expenses = Liabilities + Equity + Revenue$$

In different Financial DEA models, selected indicators represent inherent aspects that form various relevant domains and link to diverse dimensions of firm performance. For example, a Financial DEA model with assets and expenses as inputs and revenue as output measures a different relevant domain to a model with liabilities and equity as inputs and profit as output. The indicators in the two models are not interchangeable. Therefore, different accounting elements are viewed as formative indicators for the construct of firm performance.

This proposition examines the degree of the construct validity of Financial DEA when used as a benchmarking tool for firm performance. The Financial DEA efficiency scores or rankings are the measures of the construct of firm performance. As discussed in section 3.3.3, construct validity is defined as the degree to which a measure captures its theoretical construct (Cronbach & Meehl, 1955; Edwards, 2003; Nunnally & Bernstein, 1994). Within the construct validity, the convergence of measures intended to represent the same construct is convergent validity (Campbell, 1960; Campbell & Fiske, 1959; Edwards, 2003). By

contrast, divergence among measures designed to represent different constructs is discriminant validity (Campbell, 1960; Campbell & Fiske, 1959; Edwards, 2003).

Two propositions have been developed to investigate the impact of alternative accounting variables on construct validity in terms of convergent and discriminant validity (section 3.3.3). These propositions will be used in Chapter 5 to test the potential effect of alternative accounting variables on the convergent and discriminant validity of Financial DEA results.

> ***P_{3a}.*** *For the same set of indicators, alternative accounting variables generate the same constructs of firm performance showing convergent validity.*
>
> ***P_{3b}.*** *For different sets of indicators, alternative accounting variables generate different constructs of firm performance showing discriminant validity.*

The results of testing these propositions will be reported in section 5.4.

4.5. Chapter Summary

This chapter conceptualised a foundation for Financial DEA research by first reviewing 248 DEA models that exclusively incorporate financial accounting variables. As a result, 12 dimensional constructs have been identified, measuring productive efficiency (reflective model between the variable and dimensional construct), or benchmarking financial efficiency (formative model between the variable and dimensional construct). Nine indicators represented the existence of these dimensional constructs. Further, the breadth of firm performance was represented by the 12 dimensional constructs. From reflective modelling, firm performance is a general phenomenon with implicit meaning measured by dimensional constructs interchangeably. From formative modelling, firm performance is a combination of separate dimensional constructs that can be measured using different dimensional constructs.

This chapter also reviewed the methodological issues in Financial DEA application regarding the models from Financial DEA literature. Both traditional issues in DEA application and the measurement errors in measurement models were discussed.

Based on the methodological issues and the Financial DEA typology, this chapter also developed three sets of propositions to be empirically tested. The next chapter describes the results of these empirical tests and the potential impact on Financial DEA results.

Chapter 5: Phase II – Empirical Examination

5.1. Chapter Introduction

This chapter empirically examines selective measurement errors proposed in Chapter 4 (section 4.4) with simulated and archival data. The measurement errors are exemplified with selective research settings. Within each research setting, the measurement errors are quantified to examine the magnitudes of impact on the results of Financial DEA. This section is structured as below: section 5.2 examines the impact of the variation of the price factor on the results of Financial DEA. Section 5.3 examines the impact of alternative stock and flow forms of accounting variables on the results of Financial DEA. Section 5.4 examines the impact of alternative accounting variables on the results of Financial DEA. Section 5.5 summarises the chapter.

5.2. Test One: Price Factor

As stated in section 4.4.1, this test examines the impact of price variations on Financial DEA results. The test focuses on price-based aggregated accounting measures and compares their use to physical measures. This test focuses on the measurement errors introduced by price variation into the accounting measures in Financial DEA. According to conventional DEA definitions, the underlying technical efficiency is measured using physical measures of inputs and outputs. This test aims to investigate, the magnitude of impact on the efficiency estimated by Financial DEA when prices vary to various extents.

As stated in section 4.4.1, the proposition to be tested is

> ***P*$_1$**. *As variation in prices in the accounting data increase, the Financial DEA results provide less convergent validity in estimating physical technical efficiency.*

The basis for the test is that accounting variables represent the product of physical measures and prices. For example, wage expense (accounting variable) represents the number of labour hours (physical measure) multiplied by the cost per labour hour (price). Although accounting

variables are also influenced by other factors such as accounting regulations and accounting choices, this test isolates the impact of variation in prices on Financial DEA results.

This test starts by using the lens of reflective modelling, which assumes there is an underlying production process for Financial DEA. Financial DEA uses aggregated accounting data as proxies for physical measures to compute "technical efficiency". The factors of production (e.g., labour and capital) are proxied by accounting variables rather than physical variables. As discussed in section 4.4.1, researchers have discussed conceptually how the accounting data, as the product of prices and physical data, can bias the results of technical efficiency (Färe et al., 2017; Farrell, 1957; Portela., 2014). The purpose of this test is to empirically examine the impact of the price variation on Financial DEA models using simulated data.

5.2.1. Simulation Analysis

5.2.1.1. Simulation parameters

In this simulation, the Financial DEA model was conceived as reflecting an underlying physical production process. Therefore, the underlying physical efficiency is assumed to be the "true" efficiency in the simulation, which serves as a comparison point for the other scenarios.

The physical production process is used as the comparison point for the Financial DEA models calculated using derived accounting measures. To simulate the physical process, a Cobb-Douglas function was developed with two inputs (X_1, X_2) and one output (Y) (Banker & Chang, 2006; Jradi & Ruggiero, 2019).

$$Y = X_1^{0.5} X_2^{0.5}$$

As summarised in Table 5 - 1, the inputs (X_1, X_2) were generated from a normal distribution with a mean of 7.5 and a standard deviation of 0.75 (Harrison et al., 2012). The output (Y) was calculated using the Cobb-Douglas function based on the simulated inputs. Inefficiencies were generated from a half-normal distribution with a mean of 0 and a standard deviation of 0.2 (Harrison et al., 2012; Jradi & Ruggiero, 2019). These parameters were used to generate the inefficiency adjusted output (Y_{adj}). The inefficiency values were also used to calculate the "true" efficiency scores and compared with the Financial DEA results in the different price

scenarios. The same set of physical values of the two inputs and adjusted output were used in all scenarios and iterations. For the narrow and broad scenarios, different assumptions were made regarding the price factors, and the sample size was varied to examine the impact of variation in prices (Table 5 - 1).

In the constant scenario, the market prices for each input and output were constant. This scenario is an ideal case where the price for one resource (product) is the same, which is the condition that price-based aggregated data can estimate technical efficiency where all measures are physical (Färe et al., 2017). This phenomenon was simulated in the test using a coefficient of variation of 0 to generate no variation in the input and output prices. The price of input one ($W_{x_{1c}}$) was designed to be 10 for all DMUs. The price of input two ($W_{x_{2c}}$) was designed to be 15 for all DMUs. The price of the output (W_{y_c}) was designed to be 30 for all DMUs. The inputs and outputs of Financial DEA were the products of prices with corresponding physical measures calculated in the physical scenario.

In the narrow scenario, the market prices for each input and output were varied approximating a narrow distribution to simulate a market with narrow price variation. For example, in a market where firms have relatively low buying and selling power, the prices of inputs to the production process and outputs sold in the market are likely to be similar (Porter, 2011). This phenomenon was simulated using a coefficient of variation of 0.05 to generate a narrow variation in the input and output prices. The price of input one ($W_{x_{1n}}$) was generated from a normal distribution with a mean of 10 and a standard deviation of 0.5. The price of input two ($W_{x_{2n}}$) was generated from a normal distribution with a mean of 15 and a standard deviation of 0.75. The price of the output (W_{y_n}) was generated from a normal distribution with a mean of 30 and a standard deviation of 1.5. The inputs and outputs of the Financial DEA models were the products of prices with corresponding physical measures used in the physical scenario.

Table 5 - 1 Test One Parameter Specification

Scenarios	Input	Factor of interest -- Price Distribution	Output
Physical	$X_1, X_2 \sim N(7.5, 0.75)$	-	Y_{adj}
Constant	$X_1 W_{x_{1c}}, X_2 W_{x_{2c}}$	$W_{x_{1c}} = 10;\ W_{x_{2c}} = 15;\ W_{yc} = 30$	$Y_{adj} W_{yc}$
Narrow	$X_1 W_{x_{1n}}, X_2 W_{x_{2n}}$	$W_{x_{1n}} \sim N\ (10, 0.5);\ W_{x_2 n} \sim N\ (15, 0.75);\ W_{yn} \sim N\ (30, 1.5)$	$Y_{adj} W_{yn}$
Broad	$X_1 W_{x_{1b}}, X_2 W_{x_{2b}}$	$W_{x_{1b}} \sim N\ (10, 2.5);\ W_{x_{2b}} \sim N\ (15, 3.75);\ W_{yb} \sim N\ (30, 7.5) \in R^+$	$Y_{adj} W_{yb}$

Note: X_1, the first physical input. X_2, the second physical input. Y, the physical output, calculated as $Y = X_1^{0.5} X_2^{0.5}$. Y_{adj}, the adjusted physical output, with inefficiencies. $W_{x_{1c}}$, the price of the first input in the constant price scenario. $W_{x_{2c}}$, the price of the second input in the constant price scenario. W_{yc}, the price of the output in the constant price scenario. $W_{x_{1n}}$, the price of the first input in the narrow price scenario. $W_{x_2 n}$, the price of the second input in the narrow price scenario. W_{yn}, the price of the output in the narrow price scenario. $W_{x_{1b}}$, the price of the first input in the broad price scenario. $W_{x_{2b}}$, the price of the second input in the broad price scenario. W_{yb}, the price of the output in the broad price scenario.

In the broad scenario, market prices varied around broad distributions to simulate a broad price variation market. For instance, in a market where firms have strong buying and selling power, a product can be priced quite differently by different firms (Porter, 2011). Therefore, the test designed the coefficient of variation to be five times that of the narrow scenario. Mathematically, the coefficient of variation in the broad scenario was 0.25. The price of input one ($W_{x_{1b}}$) was generated from a normal distribution with a mean of 10 and a standard deviation of 2.5. The price of input two ($W_{x_{2b}}$) was generated from a normal distribution with a mean of 15 and a standard deviation of 3.75. The price of the output (W_{yb}) was generated from a normal distribution with a mean of 30 and a standard deviation of 7.5. The inputs and outputs of Financial DEA were the products of prices and the corresponding physical measures.

This test also examined whether the impact of price variation affects the Financial DEA results to different extents for different sample sizes. Four sample sizes were tested from six to 384, repeating 100 times using the Monte Carlo method (Jradi & Ruggiero, 2019). The smallest sample size was chosen to be six according to the "rule of thumb" that the number of DMUs should be no less than $2 \times (m + s)$, where m is the number of inputs and s is the number of outputs (Golany & Roll, 1989; Khezrimotlagh et al., 2019). The minimum number of DMUs under this rule is six ($2 \times (2 + 1)$).

The Financial DEA models applied a Constant Return to Scale (CRS) production function (Charnes et al.,1978) because the sum of the exponents in the Cobb-Douglas function equals one (Oh & Shin, 2015). In addition, an output orientation was chosen to match the Cobb-Douglas function. That is, the increase (decrease) of inputs will lead to a proportional increase (decrease) of output (Andor and Hesse, 2014) (section 2.2.2). To avoid any randomisation effects, each scenario was repeated 100 times (Jradi & Ruggiero, 2019; Ruggiero, 2007). There was no archival example for test one because firms usually treat price data as confidential information and do not disclose prices to external databases.

5.2.1.2. ***Simulation results***

Descriptive statistics – variables

Table 5 - 2 summarises the descriptive statistics for the simulated variables when the sample size is six and 384. The results for the sample size of 24 and 96 are reported in Appendix 2. The features of the simulated data are twofold. First, in each panel, the price factor ranges from no variation (constant scenario) to broad variation (broad scenario). Influenced by the price factor, the accounting variables also vary to different degrees. Hence, the Monte Carlo simulation is expected to show different levels of impact on the Financial DEA results associated with different variations in prices.

Second, the distributions of the variables in the same scenario are stable across the different sample sizes, in line with the simulation parameters. For example, in the physical scenario, the descriptive statistics of inputs (X_1, X_2) have a mean of approximately 0.75 and a standard deviation of approximately 7.5 across different sample sizes.

Table 5 - 2 Test One Descriptive Statistics of Simulated Data

		MEAN	SD	CV	MIN	Q1	MED	Q3	Max
Panel A: N = 6, iteration = 100									
Physical	X_1	7.48	0.75	0.10	5.11	6.98	7.47	8.01	9.61
	X_2	7.48	0.77	0.10	5.03	6.95	7.48	7.98	10.10
	Y	7.46	0.54	0.07	5.77	7.12	7.47	7.85	9.01
	Y_{adj}	6.41	0.87	0.14	3.56	5.84	6.43	7.01	8.74
Constant	W_{x1c}	10.00	0.00	0.00	10.00	10.00	10.00	10.00	10.00
	W_{x2c}	15.00	0.00	0.00	15.00	15.00	15.00	15.00	15.00
	W_{vc}	30.00	0.00	0.00	30.00	30.00	30.00	30.00	30.00
	X_1W_{x1c}	74.80	7.45	0.10	51.06	69.78	74.67	80.12	96.11
	X_2W_{x2c}	112.20	11.60	0.10	75.46	104.27	112.22	119.69	151.43
	$Y_{adj}W_{vc}$	192.23	25.97	0.14	106.74	175.24	192.87	210.30	262.14
Narrow	W_{x1n}	10.02	0.50	0.05	8.28	9.66	10.02	10.39	11.76
	W_{x2n}	14.98	0.76	0.05	12.89	14.45	14.96	15.53	17.25
	W_{vc}	30.04	1.44	0.05	25.98	29.02	30.07	31.13	34.29
	X_1W_{x1n}	74.98	8.55	0.11	50.80	69.27	74.39	80.87	100.16
	X_2W_{x2n}	112.01	12.77	0.11	74.21	103.35	111.77	120.65	156.05
	$Y_{adj}W_{vn}$	205.43	27.83	0.14	106.62	187.35	207.70	224.68	277.18
Broad	W_{x1b}	10.10	2.53	0.25	3.99	8.40	10.18	11.64	17.62
	W_{x2b}	14.98	3.71	0.25	3.67	12.61	14.89	17.43	27.79
	W_{vb}	29.74	7.55	0.25	7.65	25.15	29.46	34.59	50.57
	X_1W_{x1b}	75.55	20.54	0.27	24.57	61.10	75.29	87.96	146.96
	X_2W_{x2b}	112.34	31.08	0.28	32.61	91.22	110.31	132.92	204.07
	$Y_{adj}W_{vb}$	203.20	57.73	0.28	46.91	163.14	200.05	242.04	380.11
Panel B: N = 384, iteration = 100									
		MEAN	SD	CV	MIN	Q1	MED	Q3	Max
Physical	X_1	7.50	0.75	0.10	4.37	7.00	7.51	8.01	10.66
	X_2	7.50	0.75	0.10	3.77	7.00	7.50	8.01	10.86
	Y	7.48	0.53	0.07	5.17	7.13	7.49	7.84	9.67
	Y_{adj}	6.43	0.87	0.14	3.01	5.84	6.48	7.06	9.38
Constant	W_{x1c}	10.00	0.00	0.00	10.00	10.00	10.00	10.00	10.00
	W_{x2c}	15.00	0.00	0.00	15.00	15.00	15.00	15.00	15.00
	W_{vc}	30.00	0.00	0.00	30.00	30.00	30.00	30.00	30.00
	X_1W_{x1c}	75.02	7.50	0.10	43.69	69.99	75.06	80.11	106.58
	X_2W_{x2c}	112.56	11.25	0.10	56.54	105.01	112.57	120.19	162.97
	$Y_{adj}W_{vc}$	192.78	26.15	0.14	90.41	175.18	194.39	211.68	281.29
Narrow	W_{x1n}	10.00	0.50	0.05	7.95	9.66	10.00	10.34	12.02
	W_{x2n}	15.00	0.75	0.05	11.86	14.50	15.00	15.51	17.94
	W_{vc}	29.99	1.50	0.05	23.56	28.98	29.97	31.00	36.21
	X_1W_{x1n}	75.04	8.39	0.11	41.56	69.34	74.88	80.62	111.43
	X_2W_{x2n}	112.58	12.63	0.11	61.17	103.98	112.25	120.88	166.52
	$Y_{adj}W_{vn}$	193.58	27.93	0.14	84.68	174.61	194.42	213.15	303.43
Broad	W_{x1b}	9.99	2.49	0.25	0.10	8.32	10.00	11.68	20.03
	W_{x2b}	15.02	3.74	0.25	0.10	12.50	14.99	17.55	30.62
	W_{vb}	30.01	7.54	0.25	0.10	24.94	30.05	35.09	65.38
	X_1W_{x1b}	74.96	20.20	0.27	0.71	61.00	74.39	88.16	164.72
	X_2W_{x2b}	112.70	30.40	0.27	0.67	91.85	111.44	132.60	256.57
	$Y_{adj}W_{vb}$	193.72	55.60	0.29	0.59	155.15	191.54	229.27	468.48

Note: (a) X_1, the first physical input. X_2, the second physical input. Y, the physical output, calculated as $Y = X_1^{0.5}X_2^{0.5}$. Y_{adj}, the adjusted physical output, with inefficiencies. $W_{x_{1c}}$, the price of the first input in the constant price scenario. $W_{x_{2c}}$, the price of the second input in the constant price scenario. W_{yc}, the price of the output in the constant price scenario. $W_{x_{1n}}$, the price of the first input in the narrow price scenario. W_{x_2n}, the price of the second input in the narrow price scenario. W_{yn}, the price of the output in the narrow price scenario. $W_{x_{1b}}$, the price of the first input in the broad price scenario. $W_{x_{2b}}$, the price of the second input in the broad price scenario. W_{yb}, the price of the output in the broad price scenario. (b) The descriptive features include the minimum (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), maximum (MAX), mean (MEAN), the standard deviation (SD), and the coefficient of variation, that is the ratio of the standard deviation to the mean (CV). (c) The sample sizes are N = 6, 24, 96 ,384. The results for N = 24 and 96 are reported in Appendix 2.

Efficiency scores

Table 5 - 3 reports the efficiency scores when the sample size is six and 384. The results for the sample size of 24 and 96 are reported in Appendix 3. Overall, the physical and constant scenarios have identical results across all sample sizes. The other two scenarios, especially the broad scenario, provide different efficiency scores. The key findings are threefold.

First, the physical scenario and the constant scenario have identical results regardless of sample size. When the price factors are constant, the Financial DEA results (constant scenario) are the same as the conventional DEA results calculated using physical measures (physical scenario). As illustrated in Figure 5 - 1, the box plots for the physical and constant scenarios are distributed identically. This result can be explained by the minimum extrapolation axiom of DEA (section 2.2.2.4), which states that the frontier is estimated by the smallest function that envelops all observed DMUs (Banker et al., 1984). When a constant value changes all values, the shape of the frontier and relative ranking of the DMUs do not change.

Table 5 - 3 Test One Simulated Efficiency Scores

Panel A: N = 6, iteration = 100										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Physical	0.9168	2.07	2.25	0.4748	0.5252	0.8572	0.9500	1.0000	0.0967	0.1054
Constant	0.9168	2.07	2.25	0.4748	0.5252	0.8572	0.9500	1.0000	0.0967	0.1054
Narrow	0.9049	1.96	2.05	0.5218	0.4782	0.8376	0.9405	1.0000	0.1063	0.1175
Broad	0.7779	1.79	1.85	0.8198	0.1802	0.6044	0.8201	1.0000	0.2195	0.2822
Panel B: N = 384, iteration = 100										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Physical	0.8624	7.05	22.81	0.5718	0.4282	0.7982	0.8783	0.9427	0.0983	0.1140
Constant	0.8624	7.05	22.81	0.5718	0.4282	0.7982	0.8783	0.9427	0.0983	0.1140
Narrow	0.7790	4.10	5.51	0.6047	0.3953	0.7094	0.7861	0.8535	0.1039	0.1334
Broad	0.4349	3.39	3.60	0.9989	0.0011	0.3180	0.4129	0.5255	0.1654	0.3804

Note: (a) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs that are at least 99% efficient (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (b) All efficiency scores are generated from the CRS model. (3) The sample sizes are N = 6, 24, 96, 384. The results of N = 24 and 96 are reported in Appendix 3.

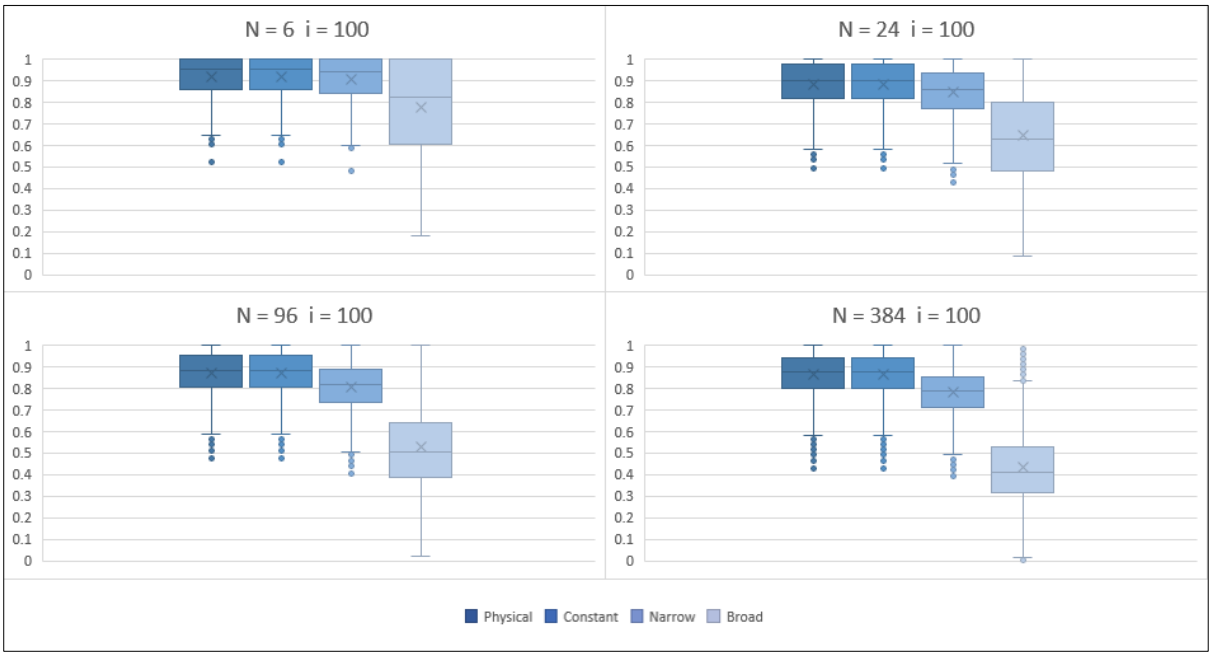


Figure 5 - 1 Test One Box Plot for the Distribution of Simulated Efficiency Scores

Second, the distribution of the Financial DEA results is influenced by the variation in prices. When the price variation is narrow, the distribution of efficiency scores is closer to the physical and constant scenarios than the broad scenario. This result is evident in Figure 5 -1 and Table 5 - 3. For example, in Table 5 - 3, Panel A (N = 6) means the efficiency scores in the physical, constant, and narrow scenarios are approximately 0.90. However, the broad scenario has a mean of 0.7779. Also, the physical, constant, and narrow scenarios have approximately 2 DMUs fully efficient, but the broad scenario only has 1.79. Further, the range of efficiency scores is approximately 0.50 in the physical, constant, and narrow scenarios. However, in the broad scenario, the range is 0.8198.

In businesses where DMUs face relatively wide variations in prices, the Financial DEA results are likely to be materially different to technical efficiency estimated using physical measures of the productive process. By comparison, when DMUs face relatively narrow prices, the Financial DEA results are likely to be similar to technical efficiency estimated using physical measures of the productive process.

Third, when the sample size increases, the discriminatory power of the Financial DEA models increases, as evidenced by the means of both the narrow and broad scenarios having reduced. For the narrow scenario, the mean reduces from 0.9049 to 0.7790 (from Panel A to B Table 5 - 3). For the broad scenario, the mean reduces from 0.7779 to 0.4349 (from Panel A to B in Table 5 - 3).

By comparison, the means in the constant and physical scenarios are relatively stable. From Panel A to B, the means drops slightly from 0.9168 to 0.8624. This finding is consistent with the previous studies whereby sample size increases are associated with systematic decreases in the mean efficiency scores (Banker et al., 1993; Diewert & Mendoza, 1995).

The decrease in efficiency in the Financial DEA results (the narrow and broad scenarios) reflects the price variations. Further, the reduction of mean efficiency scores in the broad scenario is much more significant than in the narrow scenario (Table 5 - 3). For example, in the narrow scenario, when the sample size increases from six (Panel A) to 384 (Panel B), the mean efficiency score decreases from 0.9049 to 0.7790 (by 0.1259). By comparison, in the broad scenario, the mean efficiency score decreases from 0.7779 to 0.4349 (by 0.3430).

Also, Figure 5 - 1 shows that the box plot for the physical and constant scenarios tends to be quite similar across the four panels regardless of the different sample sizes. However, the box

plots for the narrow and broad scenarios illustrate that as sample size increases, mean efficiency scores decline.

Variation in Financial DEA results

Table 5 - 4 presents the comparison of the Financial DEA efficiency results (constant, narrow, and broad) with the efficiency results calculated under the physical scenario. It shows the results for the four different sample sizes when comparing the MAD and correlation performance criteria (section 3.3.3.3). These criteria show that the constant scenario is not significantly different from the physical scenario, regardless of sample size.

First, when the price factor is constant, the results of the Financial DEA (constant scenario) is the same as the DEA (physical scenario). The estimated efficiency scores by Financial DEA and the "true" underlying efficiency scores are near identical when the price factor is constant. The physical and constant scenarios across the different sample sizes provide identical MAD, Pearson correlation, and Spearman's correlation.

Table 5 - 4 Test One Variation in Financial DEA Results

Panel A: N = 6, iteration = 100				**Panel C: N = 96, iteration = 100**			
Scenario	**MAD**	**Pearson**	**Spearman's**	**Scenario**	**MAD**	**Pearson**	**Spearman's**
Constant	0.0000	1.0000	1.0000	**Constant**	0.0000	1.0000	1.0000
		0.0000***	0.0004***			0.0000***	0.0000***
Narrow	0.0374	0.8560	0.7804	**Narrow**	0.0695	0.8655	0.8439
		0.0588*	0.1130			0.0000***	0.0000***
Broad	0.1876	0.2221	0.1830	**Broad**	0.3502	0.3096	0.3101
		0.4942	0.5021			0.0205**	0.0232**
Panel B: N = 24, iteration = 100				**Panel D: N = 384, iteration = 100**			
Scenario	**MAD**	**Pearson**	**Spearman's**	**Scenario**	**MAD**	**Pearson**	**Spearman's**
Constant	0.0000	1.0000	1.0000	**Constant**	0.0000	1.0000	1.0000
		0.0000***	0.0000***			0.0000***	0.0000***
Narrow	0.0533	0.8678	0.8273	**Narrow**	0.0863	0.8680	0.8512
		0.0000***	0.0001***			0.0000***	0.0000***
Broad	0.2645	0.2951	0.2798	**Broad**	0.4303	0.3050	0.3099
		0.2391	0.2676			0.0000***	0.0000***

Note: (a) DEA models are the constant return of scale. (b) *** for significance level of < 0.01, ** for significance level of < 0.05, * for significance level of < 0.1. (c) the criteria for the impact on Financial DEA results are the Pearson Correlation (Pearson), the Spearman's Ranking Correlation (Spearman's) and the Mean Absolute Deviation (MAD).

Second, when the price factor variation increases, the impact on the Financial DEA results increases. This test uses a numerical cut-off of 0.8 (Andor & Hesse, 2011) of the correlation coefficients to determine the degree of similarity between the test and comparison point. In addition, cut-offs of statistically significant levels of 0.01, 0.05, and 0.1 (Jradi et al., 2021) are applied. Within each panel in Table 5 - 4, the correlation coefficients of broad scenarios are below these cut-offs, regardless of sample size. For example, in Panel A and D, the Pearson correlation (Spearman's correlation) of the broad scenarios is between 0.2221 and 0.3050 (between 0.1830 and 0.3099). These correlations are not statistically significant, except where sample sizes are 96 and 384; the correlation coefficients are low but significant at 0.0000. By comparison, the narrow scenario has quite high Pearson (Spearman's) coefficients at between 0.8560 to 0.8680 (0.7804 to 0.8512), and with a higher degree of significance, the p-value is between 0.0588 and 0.0000 (0.1130 and 0.0000).

Third, when the sample size increases, the correlation between Financial DEA and physical scenario improves. However, the level of improvement depends on the variation in the level of prices. In the narrow scenario, when the sample size increases, the impact on the similarity of the Financial DEA models reduces steadily. For instance, both Pearson and Spearman's correlation coefficients are quite high (over 0.8) when the sample size increases to 24 and above, with statistically significant variances (p-values < 0.01). However, in the broad scenario, when the sample size increases, the correlation between Financial DEA and physical scenario improves, but only to a limit level. For example, although the correlations are statistically significant at the 0.01 level when the sample size is 384, the correlation coefficients remain relatively low (all below 0.31).

Fourth, when the sample size increases, the MAD increases. This result suggests that when the sample size is larger, the efficiency scores generated by the Financial DEA diverge further from the physical efficiency. In Table 5 - 4, the broad scenario has the most significant increase in MAD from N = 6 (0.1876) to N = 384 (0.4303). In contrast, the narrow scenario increase is relatively small from 0.0374 (N = 6) to 0.0863 (N = 384).

This finding aligns with the previous studies that when the "measurement error" is large, MAD increases with sample size (Banker et al., 1993; Ruggiero, 2004). The simulated data for the narrow and broad price variations have relatively larger measurement errors than physical measures. In Table 5 - 2, for the narrow scenario, the standard deviations of prices range from 0.50 to approximately 1.50. For the broad scenario, the standard deviations range

from 2.50 to 7.50. The magnitudes of the standard deviations in both scenarios would be classified as large measurement errors, according to Banker (1993).

This finding is also consistent with the findings shown in the efficiency scores, shown in Table 5 - 3. Researchers have suggested a systematic reduction of the mean efficiency score occurs with a larger sample size (Banker et al., 1993; Diewert & Mendoza, 1995). However, holding all factors except for the prices the same, the broad scenario has a lower mean efficiency score than the narrow scenario, which has a lower score than the physical and constant scenarios. When comparing the MAD in the narrow and broad scenarios, the increasing MAD is mainly driven by the incremental price variation.

Additional analysis

Figure 5 - 2 graphically demonstrates the trend of change between Spearman's correlation coefficients and the price variation based on the data points in Table 5 - 5. Table 5 - 5 was computed using the narrow (coefficient of variation = 0.05) and broad scenarios (coefficient of variation = 0.25) in Table 5 - 4, N = 384, R = 100. Table 5 - 5 also reports three additional data points calculated using the coefficient of variation of 0.10, 0.15 and 0.20 with N = 384 and R = 100.

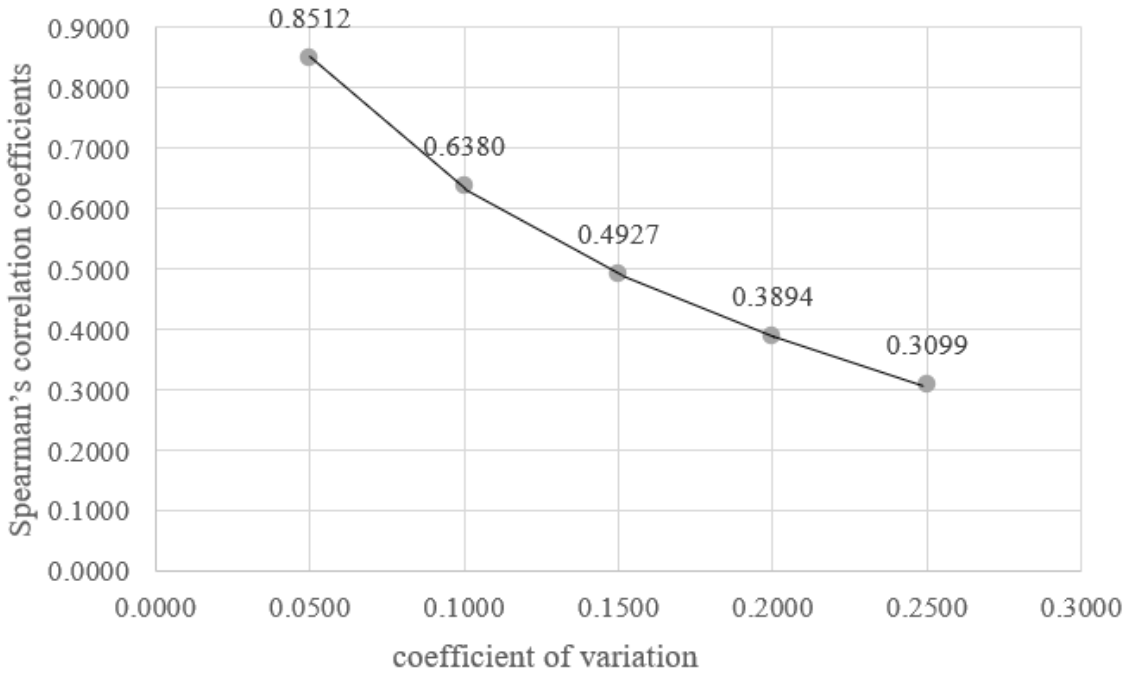


Figure 5 - 2 Test One the Trend of Change between Spearman's Correlation and Price Variation

Table 5 - 5 Test One Trend of Change

CV	Spearman's coefficient	P-value
0.0500	0.8512	0.0000***
0.1000	0.6380	0.0000***
0.1500	0.4927	0.0000***
0.2000	0.3894	0.0000***
0.2500	0.3099	0.0000***

Note: (a) *** for significance level of < 0.01, ** for significance levels of < 0.05, * for significance level of < 0.1. (b) CV is the coefficient of variation, the ratio of the standard deviation to the mean.

Figure 5 - 2 and Table 5 - 5 illustrate the relationship between price variation and its impact on Spearman's correlation between the Financial DEA results and DEA results calculated using physical measures. With the increase in price variation, correlation reduces from 0.8512 to 0.3099, although the p-values remain significant (0.0000). This additional test provides a continuous view of the impact on Financial DEA results with the change of price variation. With the increasing change of price variation, the results of Financial DEA diverge from the conventional DEA results with physical measures.

5.2.2. Summary

In summary, this section discussed the impact on the Financial DEA results where the prices vary to different degrees. The key findings are threefold. First, when the price factor is fixed, the Financial DEA results are identical to the underlying physical production process, consistent with the previous studies (Färe et al., 1990; Färe et al., 2017). Second, when the price factor varies, the greater the degree of variation introduced by prices to accounting variables, the Financial DEA results diverge more from the physical technical efficiency. Third, when the sample size increases, the Financial DEA results with little price variation improve and provide a fairly close estimation of technical efficiency. However, the Financial DEA results with wide price variation become more diverged.

5.3. Test Two: Stock and Flow Forms of Accounting Variables

This test examines how alternative forms of accounting variables act as proxies for the elements of production in Financial DEA. The test illustrates the stock and flow forms of accounting variables used to measure capital. As noted in section 4.4.2, the different forms of accounting variables measuring capital (*GPPE*, *NPPE* and *DP*) are impacted by various accounting and operational factors. The question is to what degree the choices between these options impact Financial DEA models.

As discussed in section 4.4.2, there are three propositions to be tested. The first proposition uses *DP*, compared to *GPPE*, as the comparison point, to proxy the capital used in a production process. As discussed in section 4.4.2, the interest variable to impact *DP* is useful life (*UL*). The proposition is stated as below:

> ***P*$_{2a}$.** *When the cost of the plant is fixed (GPPE), a longer estimated useful life (UL) generates smaller depreciation expense (DP) and higher Financial DEA efficiency scores when using depreciation expense (DP) as the measurement of capital.*

The second and the third propositions are using *NPPE*, comparing to *GPPE* as the comparison point, to proxy the capital used in a production process.

> ***P*$_{2b}$.** *When the cost of a plant is fixed (GPPE), and the relative age rate ($\frac{AGE}{UL}$) is the same; either useful life (UL) or age (AGE) can vary independently and calculate the other variable accordingly. As a result, if using NPPE as the proxy for capital, the Financial DEA results will be identical.*

> ***P*$_{2c}$.** *When the cost of a plant is fixed (GPPE), the higher the relative age rate ($\frac{AGE}{UL}$), the lower the NPPE and the higher the Financial DEA efficiency scores if using NPPE as a proxy for capital.*

This test starts by using the lens of reflective modelling whereby Financial DEA models are assumed to reflect underlying physical production processes. The choice of stock and flow accounting variables introduces various accounting choices and operational factors into the

measured Financial DEA results. These accounting choices and operational factors include accounting estimation, accounting calculation, age of business, and operating activities. These variations can lead to measurement errors and divergence of the Financial DEA results from the results obtained using a physical productivity DEA model. The test investigates the impact of measurement errors on the Financial DEA results due to different accounting choices and operational factors introduced by these alternative accounting measures of capital.

There are two sets of analyses in this test, with simulated data and archival data, respectively. First, a simulated test (section 5.3.1) is used to isolate the quantitative relationships between the Financial DEA results and the factors related to the accounting choices and operational variations. Second, an empirical test (section 5.3.2) is used to test whether the relationships found in the simulation are observed in real-world settings.

5.3.1. Simulation Analysis

5.3.1.1. ***Simulation parameters***

The simulated test used the known differences between the alternative accounting variables and their relationships with the underlying physical production indicators. Several assumptions had to be made about the parametric models and distributional properties to generate artificial data (Giorgio et al., 2016; Harrison et al., 2012).

The simulation followed a reflective modelling approach, which assumes the Financial DEA model reflects an underlying physical production process. To measure the production process, a Cobb-Douglas function was developed with two inputs and one output to be comparable with previous studies (Banker & Chang, 2006; Ruggiero, 2007). Different assumptions were made for the probability distribution and the sample size to test for the impact on the Financial DEA results of variations in the underlying accounting and operational factors: useful life (*UL*), age of capital (*AGE*) and relative age of capital ($\frac{AGE}{UL}$) using different accounting measures (*GPPE*, *DP* and *NPPE*) of physical capital.

This test examines the impact of variation in accounting choices and operational factors on Financial DEA results. *GPPE* was considered as the comparison point for *DP* and *NPPE* as the proxy for capital in the production process. The simulation assumed all firms are new businesses that invest in capital capacity simultaneously at the same price. The simulation

also assumed that all firms have the same buying power in the market. Therefore, the cost of *GPPE*, as the price of capital proxies the physical amount of capital. By comparison, *DP* is influenced by the accounting estimates of *UL*, and *NPPE* is influenced by both *UL* and *AGE*.

In the production process, *GPPE* is the first input, proxying the fixed capital (X_1). The second input (X_2) is operating expenses (*SIMU_OPEX*), which was assumed to proxy the sum of the labour and raw material factors of production. As summarised in Table 5 - 6, *GPPE* (*SIMU_GPPE*) and operating expense (*SIMU_OPEX*) were generated separately from normal distributions with a mean of 100 and a standard deviation of 10.[31] The output volume (Y) was sales (*SIMU_SALES*), which was generated using the Cobb-Douglas function (Banker & Chang, 2006; Ruggiero, 2007):

$$Y = X_1^{0.5} X_2^{0.5}$$

The inefficiency value was randomly generated from a half-normal distribution with a mean of zero and a standard deviation of 0.2 (Harrison et al., 2012; Ruggiero, 1999). These inefficiencies were applied to the *SIMU_SALES* to generate the adjusted sales (*SIMU_ADJ_SALES*). The value of GPPE (*SIMU_GPPE*), operating expense (*SIMU_OPEX*), and adjusted sales (*SIMU_ADJ_SALES*) were used by all Financial DEA models in the same iteration for different scenarios.

There were two factors of interest in the test: useful life (*UL*) and age (*AGE*). These two factors introduce heterogeneity into the two alternative accounting variables (*DP* and *NPPE*). To investigate the impact of the variation in these factors on the Financial DEA results, three scenarios (constant, narrow, and broad scenarios) were developed to represent different heterogeneity levels introduced by these factors.

The relative age rate determined the extent of variation in the three scenarios since, in reality, the *AGE* of a machine cannot exceed the *UL* of the machine. Therefore, the simulation generated the relative age rate first, which was set between zero and one, and based on the rate, *UL* or *AGE* were calculated for each scenario accordingly. The parameter of relative age rate was distributed around the mean of 80%. When the relative age rate is less than 50%, the

[31] The values of mean and standard deviation were adapted from the previous studies (Harrison et al., 2012), adjusted for the research context. In Harrison et al. (2012) the inputs are generated independently from normal distribution with a mean of 7.5 and a standard deviation of 0.75. That is the coefficient of variation is 0.10 (0.75/7.5). In this test, the mean of inputs is designed to be 100. Holding the coefficient of variation of 0.10, the standard deviation is 10 (0.10 × 100).

value of *NPPE* is relatively close to the *GPPE* value. As a result, the Financial DEA results calculated with *NPPE* and *GPPE* are expected to be similar.[32] In practice, firms that keep their machines in relatively new condition would reflect this relative age. By comparison, when the relative age is larger than 50%, firms tend to use the machines for a long time, and the difference between *NPPE* and *GPPE* is relatively large. This test tabulates and reports when the capital has relative *AGE* around 80% of *UL* to identify the impact of variation when *NPPE* and *GPPE* are relatively different.

To test P_{2a}, three scenarios were developed to represent different heterogeneity levels in *UL* (Table 5 - 6). In the constant scenario, *AGE_C* is 8, and *UL_C* is 10. In the narrow and broad scenarios, the factor of interest (*UL*) varied (*UL_N, UL_B*) based on the randomised relative age rates (*AGE_RATE_N, AGE_RATE_B).* Differences in *UL* affected the value of *DP* calculated based on *SIMU_GPPE*.

In the narrow scenario, the relative age rate (*AGE_RATE_N*) was first generated from a normal distribution with a mean of 0.8 and a standard deviation of 0.04 ($0.04 = 0.8 \times 0.05$). Second, based on this rate, the factor of interest (*UL_N*) was calculated as 10/*AGE_RATE_N*.[33] Third, the variable of interest (*DP_UL_N*) was calculated as *SIMU_GPPE*/ *UL_N*. For the Financial DEA model, *DP_UL_N* (the variable of interest), the *SIMU_OPEX* (other variables) were the inputs. Moreover, the output was *SIMU_ADJ_SALES* (other variables).

[32] A test was conducted using machine ages closer to zero. The relative age rate was generated from N (0.2, 0.01). Financial DEA results showed *NPPE* provided a closer estimation to *GPPE* than *DP* under this scenario. Results untabulated.

[33] The value of fixed *AGE* was designed to be 10 instead of 8; this was to ensure the *AGE* was smaller than the corresponding *UL* ($P_{2c,}$ when both *AGE* and *UL* vary). For one relative age rate, belonging to (0,1), a constant multiplied by the rate (*AGE*), is ensured to be smaller than the constant divided by the rate (*UL*). An opposite example of the given fixed *AGE* equals to 8 and the given fixed *UL* equals to 10 is illustrated as below. If the randomised relative age rate is 0.99, the *UL* calculated is 8.08 (8.08 = 8÷0.99 = given fixed *AGE*÷rate). The *AGE* calculated is 9.9 (9.9 = 10×0.99 = given fixed UL×rate). This DMU cannot exist logically since *AGE* (9.9) > *UL* (8.08).

Table 5 - 6 Test Two Parameter Specification

	Randomisation	Variation	Variable of Interest	Other Variables	Fixed
P_{2a}	*AGE_RATE_C* = 0.8 *AGE_RATE_N* *~ N (0.8, 0.04)* *AGE_RATE_B* *~ N (0.8, 0.2)*	*UL_C* = 10 *UL_N* = 10/ ***AGE_RATE_N*** *UL_B* = 10/ ***AGE_RATE_B***	*DP_UL_C = SIMU_GPPE/UL_C* *DP_UL_N = SIMU_GPPE/**UL_N*** *DP_UL_B = SIMU_GPPE/**UL_B***	*SIMU_OPEX ~ N (100, 10)* *SIMU_GPPE ~ N (100, 10)* *SIMU_SALES (equation 5 - 6)* *Inefficiency ~ N | (0, 0.2) |* *SIMU_ADJ_SALES* *= SIMU_SALES adjusted with inefficiency*	*AGE*
P_{2b} **UL**	*AGE_RATE_C* = 0.8 *AGE_RATE_N* *~ N (0.8, 0.04)* *AGE_RATE_B* *~ N (0.8, 0.2)*	*UL_C* = 10 *UL_N* =10/ ***AGE_RATE_N*** *UL_B* = 10/ ***AGE_RATE_B***	*NPPE_UL_C* *= SIMU_GPPE - DP_UL_C × AGE_C* *NPPE_UL_N* *= SIMU_GPPE - **DP_UL_N** × AGE_C* *NPPE_UL_B* *= SIMU_GPPE - **DP_UL_B** × AGE_C*	*SIMU_OPEX ~ N (100, 10)* *SIMU_GPPE ~ N (100, 10)* *SIMU_SALES (equation 5 - 6)* *Inefficiency ~ N | (0, 0.2) |* *SIMU_ADJ_SALES* *= SIMU_SALES adjusted with inefficiency*	*AGE*
P_{2b} **AGE**	*AGE_RATE_C* = 0.8 *AGE_RATE_N* *~ N (0.8, 0.04)* *AGE_RATE_B* *~ N (0.8, 0.2)*	*AGE_C* = 8 *AGE_N* = 10 × ***AGE_RATE_N*** *AGE_B* = 10 × ***AGE_RATE_B***	*NPPE_AGE_C* *= SIMU_GPPE - DP_UL_C × AGE_C* *NPPE_AGE_N* *= SIMU_GPPE - DP_UL_C × **AGE_N*** *NPPE_AGE_B* *= SIMU_GPPE - DP_UL_C × **AGE_B***	*SIMU_OPEX ~ N (100, 10)* *SIMU_GPPE ~ N (100, 10)* *SIMU_SALES (equation 5 - 6)* *Inefficiency ~ N| (0, 0.2) |* *SIMU_ADJ_SALES* *= SIMU_SALES adjusted with inefficiency*	*UL*
P_{2c}	*AGE_RATE_C* = 0.8 *AGE_RATE_N* *~ N (0.8, 0.04)* *AGE_RATE_B* *~ N (0.8, 0.2)*	*UL_C* = 10 *UL_N* = 10/ ***AGE_RATE_N*** *UL_B* = 10/ ***AGE_RATE_B*** *AGE_C* = 8 *AGE_N* = 10 × ***AGE_RATE_N*** *AGE_B* = 10 × ***AGE_RATE_B***	*NPPE_BOTH_C* *= SIMU_GPPE - DP_UL_C × AGE_C* *NPPE_BOTH_N* *= SIMU_GPPE - **DP_UL_N** × **AGE_N*** *NPPE_BOTH_B* *= SIMU_GPPE - **DP_UL_B** × **AGE_B***	*SIMU_OPEX ~ N (100, 10)* *SIMU_GPPE ~ N (100, 10)* *SIMU_SALES (equation 5 - 6)* *Inefficiency ~ N | (0, 0.2) |* *SIMU_ADJ_SALES* *= SIMU_SALES adjusted with inefficiency*	*AGE* *UL*

Note: (a) Variable definition can be found in Appendix 4. (b) For the broad scenario, a broader distribution band has been attempted, with the coefficient of variation of 2 and the standard deviation of 1.6. However, the interval of relative age rate is not significantly different from the current prescription due to the limit of (0,1). (c) The variables in bold are the results from previous steps.

In the broad scenario, first, the relative age rate (*AGE_RATE_B*) was generated from a normal distribution with a mean of 0.8 and a standard deviation of 0.2 ($0.2 = 0.8 \times 0.25$). Second, based on this rate, the factor of interest (*UL_B*) was calculated as 10/*AGE_RATE_B*. Third, the variable of interest (*DP_UL_B*) was calculated as *SIMU_GPPE*/ *UL_B*. For the Financial DEA model, the inputs were the *DP_UL_B* (the variable of interest), the *SIMU_OPEX* (other variable), and the output was *SIMU_ADJ_SALES* (other variable).

The P_{2b} test was designed using the same process as for P_{2a}. Three scenarios were designed to represent different levels of heterogeneity in *UL* (*AGE*). All parameters for P_{2b} were the same as for P_{2a}. The only difference was that in P_{2a}, the variable of interest was *DP*, and in P_{2b}, the variable of interest was *NPPE*. The value of *NPPE* was calculated as $NPPE = GPPE - AGE \times DP$.

To test P_{2c}, three scenarios were designed. However, in this test, both *UL* and *AGE* were varied so that the relative age ($\frac{AGE}{UL}$) varied. In the constant scenario, the *AGE_C* was 8, and the *UL_C* was 10.

In the narrow scenario, first, the relative life (*AGE_RATE_N*) was generated in the same way as described for P_{2a}. Second, based on this rate, the factor of interest, *UL_N* was calculated as 10 divided by the randomised relative age rate (*AGE_RATE_N*). Moreover, the other factor of interest, *AGE_N* was calculated as the product of 10 and the same randomised relative age rate (*AGE_RATE_N*). Third, based on the two factors of interest, the variable of interest (*NPPE_BOTH_N*) was calculated as *SIMU_GPPE* - *DP_UL_N* × *AGE_N*. For the Financial DEA model, the inputs were the *NPPE_BOTH_N* (other variable), the *SIMU_OPEX* (other variable), and the output was *SIMU_ADJ_SALES* (other variable).

In the broad scenario, first, the relative life (*AGE_RATE_B*) was generated in the same way as described above. Second, based on this rate, the factor of interest, *UL_B* was calculated as 10 divided by the randomised relative age rate (*AGE_RATE_B*). Furthermore, the other factor of interest, *AGE_B* was calculated as the product of 10 and the same randomised relative age rate (*AGE_RATE_B*). Third, based on the two factors of interest, the variable of interest (*NPPE_BOTH_B*) was calculated as *SIMU_GPPE* - *DP_UL_B*× *AGE_B*. For the Financial DEA model, the inputs were the *NPPE_BOTH_B* (variable of interest), the *SIMU_OPEX* (other variable), and the output was *SIMU_ADJ_SALES* (other variable).

To test the impact of sample size on the Financial DEA results, four sample sizes were tested from six to 384 and repeated 100 times using the Monte Carlo method (Andor & Hesse, 2014; Banker et al., 1993; Ruggiero, 2007). The smallest sample size was chosen to be six, according to the "rule of thumb" suggested by Golany and Roll (1989, p. 239) that "the number of units should be at least twice the number of inputs and outputs considered".

The Financial DEA models applied a Constant Return to Scale (CRS) production function (Charnes et al.,1978) because the sum of the exponents in the Cobb-Douglas function equals one (Oh & Shin, 2015). Furthermore, an output orientation was chosen to match the Cobb-Douglas function. That is, the increase (decrease) of inputs will lead to a proportional increase (decrease) of output (Andor and Hesse, 2014). Finally, to avoid any effects of randomisation, each scenario was repeated 100 times (Jradi & Ruggiero, 2019; Ruggiero, 2007).

5.3.1.2. Simulation results

Descriptive statistics – variables

Table 5 - 7 reports the descriptive statistics for the variables used in the simulation for all four sample sizes, but only N = 6 and 384 are shown in Table 5 - 7. Statistics for N = 24 and 96 are tabulated in Appendix 5. The variables in the simulation distributed align with the parameters used. The inputs (*SIMU_GPPE, SIMU_OPEX*) have a mean of approximately 100 and a standard deviation of approximately 10. The output (*SIMU_SALES*) is calculated using the Cobb-Douglas function, and the two inputs have no inefficiency. After adjusting for inefficiencies, the output (*SIMU_ADJ_SALES*) has a lower mean and a larger standard deviation than the output (*SIMU_SALES).*

The factors of interest used to calculate *DP* and *NPPE* are distributed in line with the parameters used. For example, the relative age rate (*AGE_RATE*), the absolute age (*AGE*), and the useful life (*UL*) are distributed around 0.8, 8, and 10, respectively, for the constant scenario. For each factor, the standard deviation of the broad scenario is about five times the narrow scenario, which is consistent with the research design.

Table 5 - 7 Test Two Descriptive Statistics of Simulated Data

Panel A: N = 6, iteration = 100									
	Variables	**MIN**	**Q1**	**Median**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**
	SIMU_GPPE	64.87	93.67	100.13	106.67	128.63	99.93	10.05	0.10
	SIMU_OPEX	69.86	94.11	100.46	106.88	128.84	100.22	9.58	0.10
Base	*SIMU_SALES*	81.93	95.38	99.74	104.35	117.55	99.82	6.81	0.07
	INEFFICIENCY	0.00	0.07	0.15	0.25	0.63	0.17	0.12	0.73
	SIMU_ADJ_SALES	51.14	77.49	85.74	93.52	115.26	85.20	11.36	0.13
	UL_C	10.00	10.00	10.00	10.00	10.00	10.00	0.00	0.00
	UL_N	10.37	12.05	12.49	12.94	14.82	12.52	0.65	0.05
	UL_B	10.00	10.86	12.53	15.06	39.88	13.57	3.99	0.29
	AGE_C	8.00	8.00	8.00	8.00	8.00	8.00	0.00	0.00
Factors	*AGE_N*	6.75	7.73	8.01	8.30	9.64	8.01	0.41	0.05
	AGE_B	2.51	6.64	7.98	9.21	10.00	7.82	1.68	0.21
	AGE_RATE_C	0.80	0.80	0.80	0.80	0.80	0.80	0.00	0.00
	AGE_RATE_N	0.67	0.77	0.80	0.83	0.96	0.80	0.04	0.05
	AGE_RATE_B	0.25	0.66	0.80	0.92	1.00	0.78	0.17	0.21
P_{2a}	*DP_UL_C*	6.49	9.37	10.01	10.67	12.86	9.99	1.00	0.10
DP	*DP_UL_N*	5.36	7.42	7.99	8.60	10.65	8.00	0.91	0.11
impacted by UL	*DP_UL_B*	2.54	6.54	7.83	9.31	12.20	7.82	1.87	0.24
P_{2b}	*NPPE_UL_C*	12.97	18.73	20.03	21.33	25.73	19.99	2.01	0.10
NPPE	*NPPE_UL_N*	3.97	16.71	19.92	23.05	36.13	19.92	4.56	0.23
impacted by UL or AGE	*NPPE_UL_B*	0.00	8.16	19.50	33.51	80.96	21.73	17.06	0.78
P_{2c}	*NPPE_BOTH_C*	12.97	18.73	20.03	21.33	25.73	19.99	2.01	0.10
NPPE	*NPPE_BOTH_N*	7.81	30.58	35.79	40.91	60.50	35.71	7.50	0.21
impacted by UL & AGE	*NPPE_BOTH_B*	0.00	15.68	34.98	54.90	102.68	35.91	25.37	0.71
Panel B: N = 384, iteration = 100									
	Variables	**MIN**	**Q1**	**Median**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**

	SIMU_GPPE	61.19	93.16	99.88	106.74	139.87	99.94	10.04	0.10
	SIMU_OPEX	60.51	93.27	100.04	106.81	138.66	100.04	10.02	0.10
Base	*SIMU_SALES*	70.33	94.94	99.74	104.58	129.72	99.74	7.15	0.07
	INEFFICIENCY	0.00	0.06	0.13	0.23	0.89	0.16	0.12	0.76
	SIMU_ADJ_SALES	37.73	77.84	86.46	93.91	122.42	85.61	11.62	0.14
	UL_C	10.00	10.00	10.00	10.00	10.00	10.00	0.00	0.00
	UL_N	10.54	12.09	12.49	12.94	15.57	12.53	0.63	0.05
	UL_B	10.00	10.69	12.50	15.05	49.90	13.59	4.09	0.30
	AGE_C	8.00	8.00	8.00	8.00	8.00	8.00	0.00	0.00
Factors	*AGE_N*	6.42	7.73	8.00	8.27	9.49	8.00	0.40	0.05
	AGE_B	2.00	6.65	8.00	9.36	10.00	7.84	1.72	0.22
	AGE_RATE_C	0.80	0.80	0.80	0.80	0.80	0.80	0.00	0.00
	AGE_RATE_N	0.64	0.77	0.80	0.83	0.95	0.80	0.04	0.05
	AGE_RATE_B	0.20	0.66	0.80	0.94	1.00	0.78	0.17	0.22
P_{2a}	*DP_UL_C*	6.12	9.32	9.99	10.67	13.99	9.99	1.00	0.10
DP	*DP_UL_N*	4.83	7.37	7.98	8.60	11.93	8.00	0.90	0.11
impacted by UL	*DP_UL_B*	1.74	6.51	7.92	9.26	13.42	7.83	1.90	0.24
P_{2b}	*NPPE_UL_C*	12.24	18.63	19.98	21.35	27.97	19.99	2.01	0.10
NPPE	*NPPE_UL_N*	4.70	16.86	19.81	22.92	41.26	19.99	4.50	0.23
impacted by UL or AGE	*NPPE_UL_B*	0.00	6.37	19.79	33.39	111.16	21.62	17.45	0.81
P_{2c}	*NPPE_BOTH_C*	12.24	18.63	19.98	21.35	27.97	19.99	2.01	0.10
NPPE	*NPPE_BOTH_N*	9.12	30.75	35.64	40.69	68.33	35.82	7.38	0.21
impacted by UL & AGE	*NPPE_BOTH_B*	0.00	12.33	35.54	55.52	133.98	35.60	26.01	0.73

Note: (a) P_{2a} is to test the impact of *UL* on *DP*; P_{2b} tests the impact of *UL* or *AGE* on *NPPE*; P_{2c} tests the impact of *UL* and *AGE* on *NPPE*. (2) Variable definitions can be found in Appendix 4. (b) The descriptive features include minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the maximum value (MAX), the average value (MEAN), the standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (c) The sample sizes are N = 6, 24, 96 ,384. The results of N = 24 and 96 are reported in Appendix 5.

DP has a lower standard deviation than *NPPE* when influenced by the same factor, *UL*, to the same magnitude. For instance, in the narrow scenario across the panels, the *DP_UL_N* has a standard deviation (coefficient of variation) of approximately 0.90 (0.11). By comparison, *NPPE_UL_N* has a standard deviation (coefficient of variation) of approximately 4.5 (0.23). This finding suggests that the Financial DEA results calculated using *DP* will be less impacted by variation in factors of interest than *NPPE*.

For *NPPE*, the factors of interest are *UL* or *AGE*. If the relative age rate is the same, *NPPE* is the same. The Financial DEA results using *NPPE_UL_N* (*NPPT_UL_B*) generates the same results as *NPPE_AGE_N* (*NPPE_AGE_B*) when relative age is the same. Table 5 - 7 only tabulates *NPPE* measures for *UL* (*NPPE_UL*) and both *UL* and *AGE* (*NPPE_BOTH*). When both *UL* and *AGE* vary, and the relative age is not fixed, the variation in *NPPE* is significantly greater. For example, under the narrow scenario, the standard deviation for *NPPE_UL* is approximately 4.5, and the standard deviation for *NPPE_BOTH* is approximately 7.4.

Efficiency scores

Table 5 - 8 reports the efficiency scores generated by the Financial DEA model using the simulated data for the sample N = 6 and 384. The efficiency scores generated using sample size N = 24 and 96 are reported in Appendix 6. Figure 5 - 3 depicts the box plots for the efficiency score distribution for all sample sizes. Overall, the efficiency score findings indicate that when factors of interest are constant, the Financial DEA results are identical when the capital is measured by *GPPE*, or *DP*, or *NPPE*. However, when the factors of interest vary, the Financial DEA results using *DP* are more similar to the comparison point (using *GPPE*) than the results using *NPPE*. In addition, the greater the variation of the factor, the more diverged the Financial DEA results are from the comparison point.

Table 5 - 8 Test Two Simulated Efficiency Scores

Panel A: N = 6, iteration = 100											
	Variables	**MEAN**	**100%**	**99%**	**Range**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Base	*SCORE_GPPE*	0.9194	2.08	2.35	0.4493	0.5507	0.8644	0.9549	1.0000	0.0937	0.1020
P_{2a}	*SCORE_DP_UL_C*	0.9194	2.08	2.35	0.4493	0.5507	0.8644	0.9549	1.0000	0.0937	0.1020
DP	*SCORE_DP_UL_N*	0.9167	1.99	2.25	0.4493	0.5507	0.8610	0.9497	1.0000	0.0958	0.1046
impacted by UL	*SCORE_DP_UL_B*	0.9062	2.04	2.15	0.4974	0.5026	0.8416	0.9380	1.0000	0.1041	0.1149
P_{2b}	*SCORE_NPPE_UL_C*	0.9194	2.08	2.35	0.4493	0.5507	0.8644	0.9549	1.0000	0.0937	0.1020
NPPE	*SCORE_NPPE_UL_N*	0.9099	1.97	2.14	0.4815	0.5185	0.8442	0.9357	1.0000	0.0983	0.1081
impacted by UL or AGE	*SCORE_NPPE_UL_B*	0.8932	2.00	2.18	0.4569	0.5431	0.8187	0.9213	1.0000	0.1137	0.1273
P_{2c}	*SCORE_NPPE_BOTH_C*	0.9194	2.08	2.35	0.4493	0.5507	0.8644	0.9549	1.0000	0.0937	0.1020
NPPE	*SCORE_NPPE_BOTH_N*	0.9103	1.95	2.11	0.4698	0.5302	0.8456	0.9367	1.0000	0.0977	0.1074
impacted by UL & AGE	*SCORE_NPPE_BOTH_B*	0.8930	1.99	2.18	0.4569	0.5431	0.8187	0.9206	1.0000	0.1136	0.1272
Panel B: N = 384, iteration = 100											
	Variables	**MEAN**	**100%**	**99%**	**Range**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Base	*SCORE_GPPE*	0.8621	6.87	45.52	0.5866	0.4134	0.7978	0.8783	0.9424	0.0988	0.1146
P_{2a}	*SCORE_DP_UL_C*	0.8621	6.87	30.83	0.5866	0.4134	0.7978	0.8783	0.9424	0.0988	0.1146
DP	*SCORE_DP_UL_N*	0.8347	4.78	13.04	0.5926	0.4074	0.7708	0.8486	0.9119	0.0986	0.1181
impacted by UL	*SCORE_DP_UL_B*	0.7694	3.98	27.73	0.6604	0.3396	0.7000	0.7764	0.8433	0.1044	0.1357
P_{2b}	*SCORE_NPPE_UL_C*	0.8621	6.87	27.74	0.5866	0.4134	0.7978	0.8783	0.9424	0.0988	0.1146
NPPE	*SCORE_NPPE_UL_N*	0.7740	3.93	7.92	0.6719	0.3281	0.7059	0.7812	0.8475	0.1032	0.1333
impacted by UL or AGE	*SCORE_NPPE_UL_B*	0.7509	2.24	7.92	0.6669	0.3331	0.6822	0.7559	0.8246	0.1044	0.1391
P_{2c}	*SCORE_NPPE_BOTH_C*	0.8621	6.87	7.92	0.5866	0.4134	0.7978	0.8783	0.9424	0.0988	0.1146
NPPE	*SCORE_NPPE_BOTH_N*	0.7766	3.91	7.85	0.6657	0.3343	0.7089	0.7844	0.8500	0.1027	0.1323
impacted by UL & AGE	*SCORE_NPPE_BOTH_B*	0.7508	2.23	7.83	0.6671	0.3329	0.6822	0.7559	0.8245	0.1044	0.1390

Note: (a) Variable definition can be found in Appendix 4. (b) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs that are at least 99% efficient (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (c) All efficiency scores are generated from the CRS model. (4) The sample sizes are N = 6, 24, 96,384. The results of N = 24 and 96 are reported in Appendix 6.

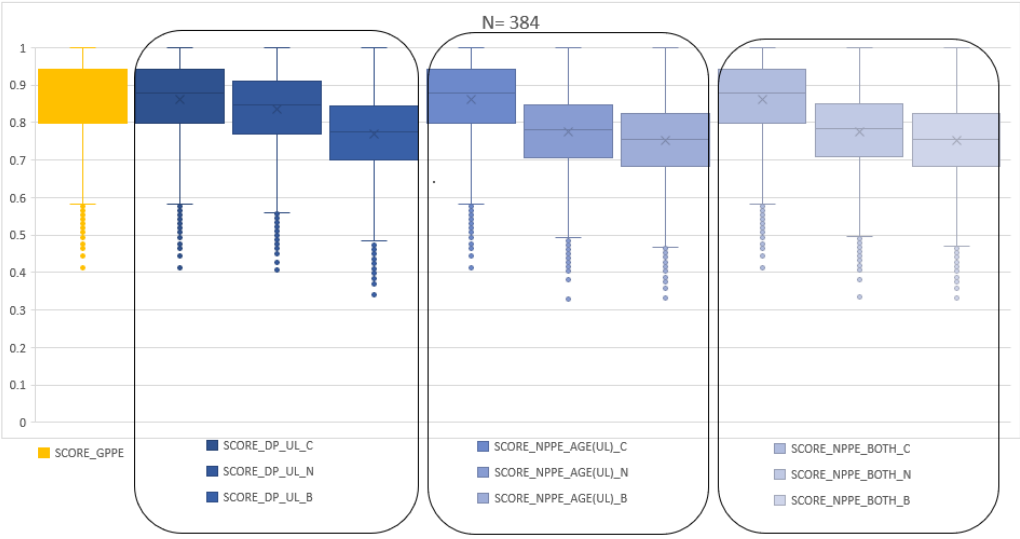


Figure 5 - 3 Test Two Box Plot for the Distribution of Simulated Efficiency Score

First, *DP* has less discriminatory power than *NPPE* as a measure of capital. For example, the mean of efficiency scores generated by Financial DEA with *DP* is higher than those generated with *NPPE*. For example, in Table 5 - 8 Panel B (N = 384), the mean of *SCORE_DP_UL_N* is 0.8347. However, the mean of *NPPE* (*SCORE_NPPE_UL_N*) is 0.7740. Further, the number of efficient DMUs identified by *DP* is higher than *NPPE*. For instance, in Table 5 - 8 Panel B (N = 384), the number of efficient DMUs (100%) from *SCORE_DP_UL_N* is 4.78. By comparison, the number of efficient DMUs from *SCORE_NPPE_UL_N* is 3.93. Similarly, the range of efficiency scores generated with *DP* is narrower than those generated with *NPPE*. In Table 5 - 8 Panel B (N = 384), the range of *SCORE_DP_UL_N* is 0.5926. However, the range of *SCORE_NPPE_UL_N* is 0.6719. These features are illustrated graphically in Figure 5 - 3.

Second, as the degree of variation in the factor of interest increases from constant to broad distribution bands, the discriminatory power of the Financial DEA increases. As a result, the efficiency scores for the DMUs calculated under the broad variation scenarios have relatively lower mean efficiency scores. For example, in Table 5 - 8 Panel B (N = 384), for *DP*, the broad scenario (*SCORE_DP_UL_B*) has a mean efficiency score of 0.7694. In contrast, the narrow scenario (*SCORE_DP_UL_N*) has a mean efficiency score of 0.8347. These features are illustrated graphically in Figure 5 - 3.

Third, when comparing across the panels, as the sample size increases from Panel A (N = 6) to Panel B (N = 384) in Table 5 - 8, the average efficiency score declines as the estimated frontier gets closer to the "true" frontier (Diewert & Mendoza, 1995). This finding aligns with the previous studies in that when the number of DMUs increases, efficient DMUs are more clearly separated from the inefficient ones (Charles et al., 2019; Dyson et al., 2001). For example, in Panel A (N = 6), for *DP*, in the narrow scenario (*SCORE_DP_UL_N*), the average efficiency score is 0.9167. However, in Panel B, for the narrow scenario of *DP* (*SCORE_DP_UL_N*), the average efficiency score drops to 0.8347. These findings are also shown graphically in Figure 5 - 3.

In summary, this section reports the descriptive statistics for the efficiency scores generated by the simulated test. The efficiency scores have the following features: (a) the efficiency scores generated by *DP* have less discriminatory power than *NPPE*. (b) The efficiency scores in the broad scenario have more discriminatory power than the narrow scenario. (c) The discriminatory power of the Financial DEA model increases with larger sample sizes.[34]

Variation in Financial DEA results

Table 5 - 9 reports on the variation in the Financial DEA results compared to the base case (*GPPE*) when the sample sizes are six and 384. The results of when sample sizes are 24 and 96 are reported in Appendix 7. Overall, the results indicate that when the factors of interest (i.e. *UL*, *AGE*, and relative age rate) are constant, the selection of alternative stock and flow forms of accounting variables (i.e. *GPPE*, *DP*, and *NPPE*) generate the same Financial DEA efficiency score. However, in reality, the factors of interest are likely to vary. Therefore, this test further investigates the impact on the Financial DEA results when a factor varies to a lesser or greater extent (i.e., the narrow and broad scenarios) when using alternative accounting measures. Several numerical cut-offs were used to assist the interpretation: (a) the correlation coefficient greater than 0.8 is recognised as a moderate level of correlation (Andor & Hesse, 2011), and the correlation coefficient greater than 0.9 is recognised as a high level of correlation (Zelenyuk, 2020). (b) The p-value of 0.01, 0.05 and 0.1 were noted for the strong, moderate, and weak level of correlation, respectively (Jradi et al., 2021).

[34] Additional tests are conducted to test the magnitude of impact of the interest factors on the corresponding Financial DEA scores. The R^2 ranges from 0.0262 to 0.0756. The finding suggests the factors of interest do not override the inefficiencies in the physical process. Results untabulated.

First, alternative accounting variables generate divergent Financial DEA results, where there is variation in the factors of interest (*UL* and/or *AGE*). When *DP* is used as the measure of capital, it has less discriminatory power than *NPPE*. Table 5 - 9 illustrates a general trend that the Financial DEA efficiency scores generated using *DP* (P_{2a}) are more closely related to the scores generated by *GPPE* (the comparison point) than *NPPE* (P_{2b} and P_{2c}). This difference is obvious when the sample size is relatively small. For example, in Table 5 - 9, Panel A, when the sample size just meets the "rule of thumb" (N = 6), only the Financial DEA results using *DP* (P_{2a}) are significantly correlated with the Financial DEA results generated using *GPPE*, at less than the 0.05 level. The Financial DEA results calculated using *NPPE* (P_{2b} and P_{2c}) are impacted by more factors (i.e. the *AGE*) and tend to have relatively lower correlation coefficients and statistically insignificant p-values.

Second, the greater the variation in the factors (*UL* and/or *AGE*), the greater the divergence of the scores of Financial DEA from the comparison point. As the degree of variation in the factor of interest increases from the constant to the broad distribution bands, the discriminatory power of the Financial DEA increases. This finding is supported by relatively lower correlation coefficients, and larger p-values, and MAD values. For instance, in Table 5 - 9 Panel A, when the *UL* impacts the Financial DEA results using *DP* (P_{2a}), the correlation and MAD in the constant scenario are almost perfectly correlated compared to the comparison point (correlation coefficients are 1, p-values are approximately 0.0000, and MAD is 0.0000). However, in the narrow scenario, correlation coefficients reduce to approximately 0.9, with less significant p-values (increase to the interval of 0.1 to 0.5). Furthermore, in the broad scenario, the correlation coefficient decreases to approximately 0.8, with barely significant p-values (approximately 0.1). These patterns can also be found in *NPPE* (P_{2b} and P_{2c}) and for different sample sizes.

Table 5 - 9 Test Two Variation in Financial DEA Results (Simulated)

Panel A: N = 6, iteration = 100									
	Constant			**Narrow**			**Broad**		
	Pearson	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**
P_{2a}	1.0000	1.0000	0.0000	0.9485	0.9101	0.0143	0.8155	0.7587	0.0360
DP impacted by UL	0.0000***	0.0005***		0.0146**	0.0301**		0.0979*	0.1346	
P_{2b}	1.0000	1.0000	0.0000	0.8149	0.7673	0.0356	0.6889	0.6343	0.0536
NPPE impacted by UL or AGE	0.0000***	0.0005***		0.0963*	0.1324		0.1801	0.2307	
P_{2c}	1.0000	1.0000	0.0000	0.8299	0.7780	0.0337	0.6914	0.6367	0.0532
NPPE impacted by UL & AGE	0.0000***	0.0005***		0.0848*	0.1224		0.1764	0.2283	
Panel B: N = 384, iteration = 100									
	Constant			**Narrow**			**Broad**		
	Pearson	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**
P_{2a}	1.0000	1.0000	0.0000	0.9743	0.9665	0.0300	0.8581	0.8389	0.0951
DP impacted by UL	0.0000***	0.0000***		0.0000***	0.0000***		0.0000***	0.0000***	
P_{2b}	1.0000	1.0000	0.0000	0.8652	0.8462	0.0906	0.8394	0.8166	0.1121
NPPE impacted by UL or AGE	0.0000***	0.0000***		0.0000***	0.0000***		0.0000***	0.0000***	
P_{2c}	1.0000	1.0000	0.0000	0.8718	0.8533	0.0880	0.8402	0.8175	0.1121
NPPE impacted by UL & AGE	0.0000***	0.0000***		0.0000***	0.0000***		0.0000***	0.0000***	

Note: (a) P_{2a} is to test the impact of *UL* on *DP*; P_{2b} tests the impact of *UL* or *AGE* on *NPPE*; P_{2c} tests the impact of *UL* and *AGE* on *NPPE*. (b) DEA models are the constant return of scale. (c) *** for significance level of < 0.01, ** for significance level of < 0.05, * for significance level of < 0.1. (d) the criteria for Financial DEA results are the Pearson Correlation (Pearson), the Spearman's Ranking Correlation (Spearman's), the Mean Absolute Deviation (MAD), the mean value of efficiency scores (MEAN), the number of efficient DMUs (100%), and the range of efficiency scores (RANGE). (e) The comparison point is the Financial DEA results generated using GPPE. (6) the sample size covers N = 6, N = 24, N = 96, N = 384. The results of N = 24, N = 96 have the same pattern as N = 6 and N = 384. The detailed table can be found in Appendix 7.

Third, when sample sizes increase across panels (N = 6 to N = 384), the correlation between the Financial DEA results using *DP* (P_{2a}) and *NPPE* (P_{2b} and P_{2c}) compared to *GPPE* improve. The mean efficiency score declines as the estimated frontier gets closer to the "true" frontier (Diewert & Mendoza, 1995). However, from the perspective of the MAD, the vertical distance between Financial DEA scores increases. In Table 5 - 9, the correlation results improve in both the narrow and broad scenarios when comparing Panel A and B. For example, in Panel A, P_{2b}, (N = 6, *NPPE* impacted by *UL* or *AGE*), the Pearson (Spearman's) correlation coefficient is 0.8149 (0.7673) with p-values of 0.0963 (0.1324). In Panel B, P_{2b}, (N = 384, *NPPE* impacted by *UL* or *AGE*), the Pearson (Spearman's) correlation coefficients improved to 0.8652 (0.8462) with highly significant p-values at 0.0000 (0.0000). This trend can also be found in P_{2a} (*DP*) and across other sample sizes reported in Appendix 7. This finding indicates that when the sample size is relatively large, the correlations improve between the Financial DEA results using *DP* (P_{2a}) and *NPPE* (P_{2c} and P_{2c}) compared to *GPPE.* Although the results using *DP* consistently generated the highest correlations with the comparison point from the small to large sample sizes, when the sample size is N = 364, the results using the *GPPE*, *NPPE* or *DP* are relatively similar in terms of rankings.

Fourth, MAD increases with sample size. This result suggests that when the sample size is relatively large, the efficiency scores generated by Financial DEA with *DP* (P_{2a}) and *NPPE* (P_{2b} and P_{2c}) diverge further from the comparison point (*GPPE*). This finding aligns with the previous studies that when measurement error is large, MAD increases with sample size (Banker et al., 1993; Ruggiero, 2004). In this test, the simulated data have relatively large "measurement errors" compared to the comparison point. In Table 5 - 7, standard deviations of the variable of interest range from 2 to 20, which are classified as larger measurement errors, according to Banker (1993).

5.3.2. Empirical Analysis

This section reports the data (section 5.3.2.1), models (section 5.3.2.2), and results (section 5.3.2.3) of the empirical test. The empirical test aims to examine whether the relationships found in the simulated test are realistic when using archival data. As noted, one of the weaknesses of simulation is that the assumptions made may be unrealistic. For example, in the simulated test, the relative life is assumed to have a mean of 80%, which can be unrealistic if an industry is relatively new. Also, the simulated test assumed a straight-line

depreciation method. In practice, firms choose the depreciation method depends on industry and machine features.

5.3.2.1. Empirical data

The simulated test used artificial data based on a set of assumptions, parametric functions, and distributional properties, which may limit the generalisability of the results whereby the artificial data does not represent reality (Harrison et al., 2012). As a remedy, an empirical test was also conducted to examine whether similar results to the simulations were observed. It is expected that the relationships found in the simulation will generally reflect the same trends as the empirical data.

The empirical data were retrieved from the COMPUSTAT database, covering five years of financial data to provide a large enough sample size without being impacted by significant technology change (Färe et al., 1994; Golany & Roll, 1989).

The sample covered two industries, the box (SIC 2440 – 3412 Shipping Containers) and the gold (SIC 1040 – 1049 Precious Metals) industries according to the Fama-French 48 industry classification (Demerjian et al., 2012; Fama & French, 1997). These two industries were selected to represent a more homogeneous industry (the box industry) and a more heterogeneous industry (the gold industry). These industries were chosen based on the mean efficiency scores reported by Demerjian (2018). The box industry had a mean efficiency score of 0.934, which indicated a high level of homogeneity in the industry (Demerjian, 2018). The gold industry had a mean efficiency score of 0.320, suggesting a low homogeneity level (Demerjian, 2018).

Table 5 - 10 Test Two Archival Sample Description

Industry	**Box**	**Gold**
All COMPUSTAT observations from 2015-2019	59	409
Delete observations with missing values	-2	-23
Delete observations with zero values	-0	-125
Delete unreasonable extreme values	-0	-8
Full sample	57	253

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

As reported in Table 5 -10, the box industry had an initial sample size of 59, and the gold industry had an initial sample of 409. Next, all observations with missing value and zero values were excluded (Golany & Roll, 1989). Last, the observations with extreme values

(estimated *UL* and *AGE* over 100 years, Table 5 - 11) were excluded. There were in total eight firm-years excluded from the sample set: Equinox Gold Corp 2016, Equinox Gold Corp 2017, Equinox Gold Corp 2018, the Scorpio Gold Corp 2018, Scorpio Gold Corp 2019, Dynaresource Inc 2018, Dynaresource Inc 2019, and the New Jersey Mining Co 2016. The final sample of the box industry contained 57 DMUs, and the gold industry had 253 DMUs. All values were inflated by the consumer price index (CPI) to the values of the financial year 2019 (Rouse & Tripe, 2016) to remove the impact of inflation. A sample size of 30 DMUs was resampled with replacement from the final samples in Table 5 -10 to provide comparably sized samples. To get relatively randomised results, each empirical sample set was resampled 1000 times (Efron, 1993).

5.3.2.2. ***Empirical models***

The accounting variables used in the Financial DEA models were the same as the simulated test. The first input was operating expenses (*OPEX*), measuring the labour and other non-capital expenses. The second input was capital, measured by *GPPE*, *NPPE* or *DP* alternatively. The output was sales (*SALES*).

The empirical test was analysed using constant returns to scale (CRS) and variable returns to scale (VRS). The CRS was used to maintain consistency with the simulated test. The VRS model was used since the empirical samples have a wide range of firm scale (Banker et al., 1984).

5.3.2.3. ***Empirical results***

Descriptive statistics – variables

Table 5 - 11reports the descriptive statistics of the archival data used. The variables used are the same as the simulated test, except for the *UL* and *AGE*, which are not publicly available. Instead, these variables were estimated using the same equations and assumptions used in the simulation test. The assumptions are that there are zero residual values for all PPE and that the depreciation method used is the straight-line method. As a result, the *UL* is calculated as the ratio of *GPPE* and *DP*. The *AGE* is calculated as the difference between the *GPPE* and *NPPE*, divided by *DP*. Therefore, some of the calculated values are relatively large. For example, the maximum value of *GOLD_AGE* is 81.88, and the maximum of *GOLD_UL* is

96.63 years.[35] These assumptions do not capture all accounting choices, and operational factors applied in practice.

Table 5 - 11 Test Two Descriptive Statistics of Raw Archival Data

Panel A Box Industry FY 2015 - 2019 (N = 57)								
	MIN	**Q1**	**MED**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**
BOX_GPPE	700.11	3757.26	5717.92	7253.70	17461.30	5935.88	3550.00	0.60
BOX_DP	39.05	206.48	386.53	490.00	1440.50	420.02	307.48	0.73
BOX_NPPE	275.56	1339.47	2968.90	3456.60	11189.50	3132.78	2522.21	0.81
BOX_OPEX	711.38	3883.60	5398.58	7762.38	15148.10	5991.54	3261.59	0.54
BOX_SALES	785.60	4595.00	6660.00	9219.88	18289.00	7116.49	3913.50	0.55
BOX_UL	6.72	13.37	16.22	17.94	28.01	15.85	4.26	0.27
BOX_AGE	2.61	5.88	8.70	10.43	14.89	8.34	3.19	0.38
BOX_R	0.24	0.44	0.53	0.58	0.68	0.51	0.11	0.21
Panel B Gold Industry FY 2015 - 2019 (N = 253)								
	MIN	**Q1**	**MED**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**
GOLD GPPE	0.27	111.55	739.24	4023.72	62307.00	4374.88	9198.24	2.10
GOLD DP	0.03	5.01	40.90	222.26	2240.00	209.62	390.69	1.86
GOLD NPPE	0.12	74.63	408.73	2260.54	25276.00	2005.51	3624.51	1.81
GOLD OPEX	0.60	40.39	183.69	671.78	7040.00	699.57	1221.33	1.75
GOLD SALES	0.00	54.93	231.89	866.81	9938.50	981.89	1861.62	1.90
GOLD_UL	0.01	13.88	19.52	28.07	96.93	23.65	15.95	0.67
GOLD_AGE	0.01	4.48	7.54	13.39	81.88	10.41	10.33	0.99
GOLD_R	0.02	0.27	0.45	0.62	0.93	0.45	0.23	0.51

Note: (a) All variable definitions can be found in Appendix 4. (b) All variables, except for the useful life, age, and relative age rate, are retrieved from the COMPUSTAT database and in the unit of millions United States dollars (USD). (c) The value of useful life, age, and relative age rate are estimates based on accounting definitions, assuming zero residual value and straight-line calculation method.

For example, in the annual reports of the gold industry, the units-of-activity depreciation method is a common choice. In the FY 2019 annual report of AngloGold Ashanti LTD the report stated, "the majority of mining assets are amortised using the units-of-production method where the mine operating plan calls for production from a well-defined proved and

[35] The DMUs with extreme values were excluded (i.e. the UL and AGE over 100 years). There are 8 firm-years excluded from the sample set: Equinox Gold Corp 2016, Equinox Gold Corp 2017, Equinox Gold Corp 2018, the Scorpio Gold Corp 2018, Scorpio Gold Corp 2019, Dynaresource Inc 2018, Dynaresource Inc 2019, and the New Jersey Mining Co 2016. The DUMs with unreasonable extreme values are feature by relatively large GPPE and relatively small DP. For example, Scorpio Gold Corp in FY2018 had the GPPE of $124.71 million USD, DP of $0.06 million USD and NPPE of $8.18 million USD. This leads the estimated UL of 2103.72 years and AGE of 1965.67 years.

probable Ore Reserve" (Anglo Gold Ashanti, 2019, p. 36). However, the units-of-activity depreciation method is characterised by a significant variation in activity units from year to year. Also, estimating total activity and recording actual activity can be difficult (Weygandt et al., 2015). This feature of the gold industry suggests that the Financial DEA results can be impacted by the unique feature of the accounting method captured by the accounting variables.

In general, Table 5 - 11 indicates that the box industry represents a relatively homogeneous sample set, and the gold industry represents a relatively heterogeneous sample set. The overall coefficient of variation of all the box industry variables (range from 0.2 to 0.8) is relatively smaller than the gold industry (range from 0.5 to 2.1). In the box industry, the factors of interest, the *UL*, *AGE*, and the relative age rate are relatively homogeneous, whereas they are much less homogeneous in the gold industry. The standard deviation (coefficient of variation) of *GOLD_UL* is 15.95 (0.6) in the gold industry and is approximately 4 (2.5) times larger than the corresponding value in the box industry, which has a value of 4.26 (0.2). The standard deviation (coefficient of variation) of *GOLD_AGE* is 10.33 (0.9) in the gold industry and is approximately 3 (3) times the box industry, which has a value of 3.19 (0.3). The standard deviation (coefficient of variation) of the relative age rate, *GOLD_R*, is 0.23 (0.5) in the gold industry and is approximately 2 (2) times larger than the box industry, which has the value of 0.11 (0.2).

Table 5 - 12 reports the descriptive statistics of the bootstrapped data. Overall, the resampling keeps the key features of the raw sample in Table 5 - 11. The key descriptive measures in Table 5 - 12 are similar to the ones in Table 5 - 11.

Table 5 - 12 Test Two Descriptive Statistics of Bootstrapped Data

Panel A Box Industry (N = 30, iteration = 1000)								
	MIN	**Q1**	**MED**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**
BOOT_BOX_GPPE	700.11	3757.26	5717.92	7253.70	17461.30	5939.74	3547.40	0.60
BOOT_BOX_DP	39.05	206.48	386.53	490.00	1440.50	420.21	307.09	0.73
BOOT_BOX_NPPE	275.56	1339.47	2968.90	3456.60	11189.50	3135.67	2517.22	0.80
BOOT_BOX_OPEX	711.38	3883.60	5398.58	7762.38	15148.10	5981.42	3253.53	0.54
BOOT_BOX_SALES	785.60	4595.00	6660.00	9219.88	18289.00	7106.15	3904.33	0.55
BOOT_BOX_UL	6.72	13.37	16.22	17.94	28.01	15.85	4.20	0.27
BOOT_BOX_AGE	2.61	5.88	8.70	10.43	14.89	8.35	3.16	0.38
BOOT_BOX_R	0.24	0.44	0.53	0.58	0.68	0.51	0.11	0.21
Panel B Gold Industry (N = 30, iteration = 1000)								
	MIN	**Q1**	**MED**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**
BOOT_GOLD_GPPE	0.27	110.07	739.20	4058.12	62307.00	4342.86	9083.08	2.09
BOOT_GOLD_DP	0.03	4.86	40.48	231.67	2240.00	209.39	389.37	1.86
BOOT_GOLD_NPPE	0.12	70.06	398.67	2287.40	25276.00	1997.42	3603.36	1.80
BOOT_GOLD_OPEX	0.60	38.54	183.05	667.43	7040.00	699.21	1219.66	1.74
BOOT_GOLD_SALES	0.00	49.32	227.08	867.89	9938.50	981.16	1856.25	1.89
BOOT_GOLD_UL	0.01	13.89	19.56	28.22	96.93	23.66	15.94	0.67
BOOT_GOLD_AGE	0.01	4.48	7.53	13.29	81.88	10.36	10.25	0.99
BOOT_GOLD_R	0.02	0.26	0.45	0.61	0.93	0.45	0.23	0.51

Note: (a) All variable definitions can be found in Appendix 4. (b) All variables, except for the useful life age and relative age rate, are bootstrapped based on the real-world data retrieved from the COMPUSTAT database and are in unit of millions United States dollars (USD). (c) The value of useful life, age, and relative age rate are estimates based on accounting definitions, assuming zero residual value and straight-line calculation method.

For example, the *BOX_OPEX* has a difference of 10.12 on the mean between the raw data (5991.54) and the bootstrapped data (5981.42). Additionally, the value of the coefficient of variation is identical between the raw data and the bootstrapped data for all variables. The bootstrapping solved the dimensionality issue by resampling 30 DMUs from the two industries for every iteration so that the results from the two industries are not affected by differences in sample size.

Efficiency scores

Table 5 -13 reports the efficiency scores. Overall, the VRS input-oriented (Panel A) and VRS output-oriented results (Panel B) are very similar. By comparison, the CRS results (Panel C) are less similar due to different economies of scale.

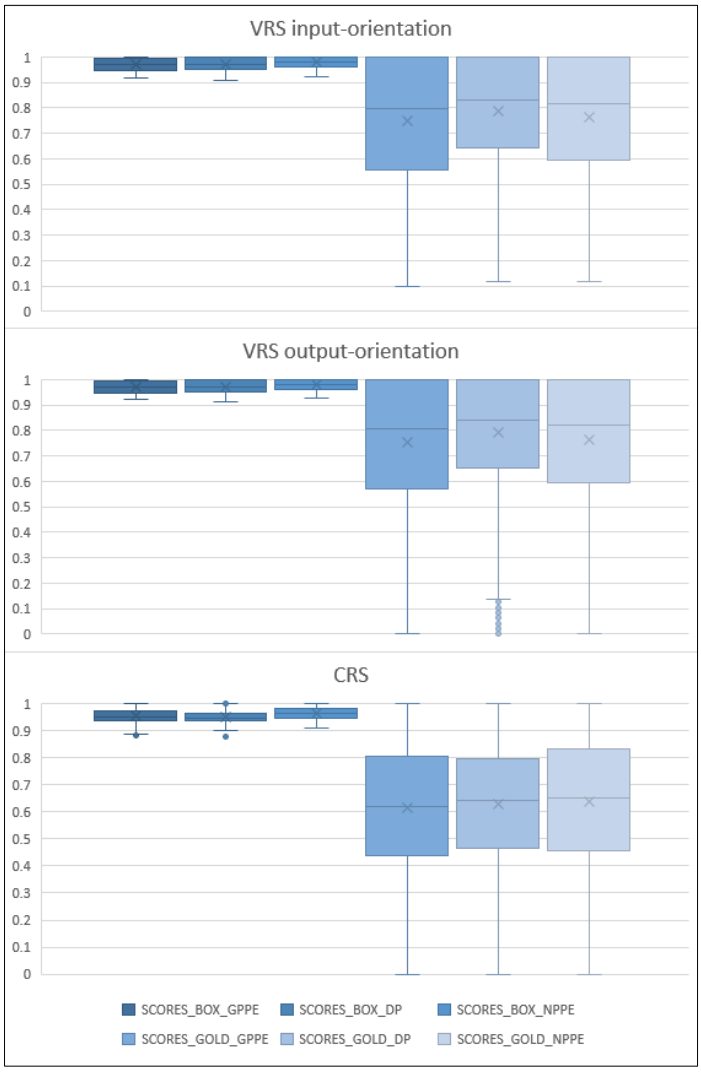


Figure 5 - 4 Test Two Box Plot for the Distribution of Bootstrapped Efficiency Scores

Table 5 - 13 Test Two Bootstrapped Efficiency Scores

Panel A: VRS Input-orientation										
	MEAN	**100%**	**99%**	**Range**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
SCORES_BOX_GPPE	0.9699	7.05	9.41	0.0828	0.9172	0.9484	0.9715	0.9968	0.0244	0.0251
SCORES_BOX_DP	0.9696	7.28	8.52	0.0920	0.9080	0.9502	0.9699	0.9993	0.0244	0.0251
SCORES_BOX_NPPE	0.9782	9.56	12.70	0.0759	0.9241	0.9605	0.9818	1.0000	0.0217	0.0222
SCORES_GOLD_GPPE	0.7495	9.05	9.36	0.9010	0.0990	0.5582	0.7965	1.0000	0.2439	0.3253
SCORES_GOLD_DP	0.7873	9.15	9.46	0.8811	0.1189	0.6422	0.8311	1.0000	0.2151	0.2732
SCORES_GOLD_NPPE	0.7646	8.66	8.98	0.8838	0.1162	0.5950	0.8150	1.0000	0.2331	0.3048
Panel B: VRS Output-orientation										
	MEAN	**100%**	**99%**	**Range**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
SCORES_BOX_GPPE	0.9694	7.05	9.32	0.0793	0.9207	0.9475	0.9703	0.9970	0.0245	0.0252
SCORES_BOX_DP	0.9700	7.28	8.52	0.0865	0.9135	0.9500	0.9701	0.9993	0.0237	0.0244
SCORES_BOX_NPPE	0.9782	9.56	12.59	0.0749	0.9251	0.9606	0.9812	1.0000	0.0216	0.0220
SCORES_GOLD_GPPE	0.7523	9.05	9.38	0.9995	0.0005	0.5703	0.8053	1.0000	0.2481	0.3299
SCORES_GOLD_DP	0.7905	9.15	9.49	0.9990	0.0010	0.6541	0.8418	1.0000	0.2222	0.2810
SCORES_GOLD_NPPE	0.7642	8.66	9.00	0.9996	0.0004	0.5968	0.8193	1.0000	0.2411	0.3155
Panel C: CRS										
	MEAN	**100%**	**99%**	**Range**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
SCORES_BOX_GPPE	0.9536	2.56	3.68	0.1194	0.8806	0.9375	0.9502	0.9724	0.0270	0.0283
SCORES_BOX_DP	0.9491	2.58	3.40	0.1205	0.8795	0.9356	0.9462	0.9613	0.0268	0.0282
SCORES_BOX_NPPE	0.9623	2.49	4.24	0.0907	0.9093	0.9462	0.9623	0.9793	0.0231	0.0240
SCORES_GOLD_GPPE	0.6159	3.23	3.38	0.9995	0.0005	0.4390	0.6181	0.8049	0.2510	0.4076
SCORES_GOLD_DP	0.6290	3.16	3.28	0.9993	0.0007	0.4656	0.6438	0.7975	0.2406	0.3826
SCORES_GOLD_NPPE	0.6357	3.39	3.55	0.9996	0.0004	0.4540	0.6512	0.8325	0.2520	0.3963

Note: (a) All variable definitions can be found in Appendix 4. (b) This table reports the key features of the Financial DEA scores with bootstrapped empirical data. For each iteration, the sample size is 30, which is repeated 1000 times. (c) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs with efficiency scores over 99% (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV).

The mean efficiencies for the box industry are all higher than for the gold industry, consistent with the expected level of homogeneity. For example, in Panel A, the box industry's mean efficiency scores are above 0.95 in all models. By comparison, the mean efficiency scores in the gold industry are approximately 0.80. Further, the standard deviation (coefficient of variation) in the box industry is approximately 0.02 (0.02), whereas, in the gold industry, it is approximately 0.23 (0.30). The standard deviation (coefficient of variation) of the gold industry is approximately 10 (12) times that of the box industry. These features are also illustrated graphically in Figure 5 - 4 by box plots, demonstrating the wide variation between efficient DMUs and the inefficient DMUs in the gold industry compared to the box industry.

Variation in Financial DEA results

Table 5 - 14 reports the variation in the Financial DEA results using the archival data. Overall, the box industry has the same relationship as the simulated test. That is, compared to the Financial DEA results using *GPPE*, *DP* generates closer results than *NPPE*. However, in the gold industry, an opposite relationship is found. Compared to the Financial DEA results using *GPPE*, *DP* generates more divergent results than *NPPE*. For instance, for the box industry, in Table 5 - 14, Panel A, input orientation, the Pearson (Spearman's) correlation coefficient of *DP* is 0.7315 (0.7245). It is higher than the one for *NPPE*, which is 0.6610 (0.6579). In this panel, the MAD of the efficiency scores generated with *DP* is 0.0118, which is slightly closer to the comparison point than the one generated by *NPPE* (0.0137).

Table 5 - 14 Test Two Variation in Financial DEA Results (Bootstrapped)

Panel A: Box 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	0.7245	0.0118	0.7333	0.7290	0.0116	0.6506	0.6057	0.0160
	0.0007***	0.0008***		0.0009***	0.0009***		0.0048***	0.0112**	
NPPE	0.6610	0.6579	0.0137	0.6527	0.6512	0.0138	0.6046	0.5721	0.0184
	0.0017***	0.0014***		0.0019***	0.0016***		0.0028***	0.0050***	
Panel B: Gold 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	0.7395	0.6914	0.1050	0.7232	0.6736	0.1066	0.8047	0.7602	0.0901
	0.0016***	0.0019***		0.0042***	0.0028***		0.0005***	0.0012***	
NPPE	0.8503	0.8116	0.0701	0.8455	0.8049	0.0711	0.8688	0.8254	0.0699
	0.0000***	0.0001***		0.0002***	0.0003***		0.0001***	0.0003***	

Note: (a) DEA models are the constant return of scale. (b) *** for significance level of < 0.01, ** for significance level of < 0.05, * for significance level of < 0.1. (c) The criteria for the variation of Financial DEA results are the Pearson Correlation (Pearson), the Spearman's Ranking Correlation (Spearman's) and the Mean Absolute Deviation (MAD). (d) The comparison point is the Financial DEA results generated using *GPPE*.

In practice, these differences could be due to specific features of the box industry (SIC 2440 – 3412 Shipping Container). The box industry provides shipping services and packing goods for customers around the world. The plant used in the box industry is mainly container boxes made of steel, which have relatively short useful lives and are updated frequently to ensure security during travelling (Wan et al., 2016). The box industry has developed in the past 50 years due to growth in international trade (Lee et al., 2014). Given the state of the international market and that companies are relatively more recently established, the box industry's *GPPE* value is likely to be relatively homogeneous. By comparison, *DP* is relatively more heterogeneous since the *UL* varies depending on the goods to be shipped and the length of the route to be travelled. Containers subject to more rough travelling environments are scrapped more frequently (Lee et al., 2014).

Further, the value of *NPPE* is the most heterogeneous variable in the box industry since the *AGE* varies in addition to the *UL* due to the length of business, the change of customer relationships, or the durability of the box's material (Ko et al., 2020). As a result, the relative homogeneity of *GPPE, NPPE* and *DP* is similar to the simulated test. That is, *GPPE* is more homogeneous than *DP* and *NPPE*.

By comparison, in the gold industry, the relationships of the Financial DEA results generated by the *NPPE* and *DP* compared to results using *GPPE* are the opposite. *NPPE* and *GPPE* generate similar Financial DEA results. By comparison, the Financial DEA results using *DP* are relatively different from the *GPPE* results. For instance, in Table 5 - 14, Panel B, input orientation, the Pearson (Spearman's) correlation coefficient of *DP* is 0.7395 (0.6914). It is lower than that of *NPPE*, which is 0.8503 (0.8116). In this panel, the MAD of the efficiency scores generated with *DP* is 0.1050, further from the comparison point than the one generated by *NPPE* (0.0701).

The gold industry (SIC 1040 – 1049 Precious Metal) has unique features that are quite different from the box industry or the simulation assumptions. As noted in Table 5 - 12, the gold industry has relatively heterogeneous *GPPE* due to its long history and the unique geology of the ore bodies mined by different firms (Lakshmanan & Gorain, 2019). Firms in the gold industry tend to vary in age (James, 1992). Gold coins were first produced in the sixth century Before the Common Era (BCE) (Lakshmanan & Gorain, 2019). During such a long history, mining spots open up when new sites are detected. Furthermore, machines used in the gold industry are purchased at very different times, depending on the age of the mines.

Also, the types of machines needed for each site depend on the site's geographical features, which adds further variation to the cost of the machine (Lakshmanan & Gorain, 2019). Therefore, the *GPPE* is relatively heterogeneous in the gold industry.

Second, the most commonly used depreciation method utilised by the gold industry is the unit-of-activity depreciation method.[36] This depreciation method is characterised by variation in the amount due to significant productivity variation from one year to another (Weygandt et al., 2015). Another feature of this method is the difficulty of estimating the total units of activity and the record of actual units of activity (Weygandt et al., 2015). As a result, in the gold industry, the DP for each year is unlikely to be a constant portion of *GPPE*, as it is in the straight-line method.

Accordingly, in the gold industry, *NPPE* can be quite similar to *GPPE* due to the large portion of non-depreciable land and assets under construction used in the industry. For instance, in the financial report of Anglo Gold Ashanti LTD 2019, "the total cost of the assets under construction and land and buildings is around 20% of the total cost of PPE" (Anglo Gold Ashanti, 2019, p. 94). Furthermore, in Table 5 - 12, the mean (median) of the relative age rate is 45% (45%) in the gold industry, compared to 51% (53%) in the box industry and 80% (80%) used in the simulation. Consequently, in the gold industry, the *NPPE*, representing the future value of PPE and the cost of non-depreciable amount, is more similar to *GPPE*.

The Financial DEA models provide more discrimination between the efficient DMUs and the inefficient DMUs in the gold industry than in the box industry. For example, in Table 5 - 14, when comparing across panels, the difference between the correlation coefficients in the gold industry are greater than the ones in the box industry. For example, for the Pearson correlation coefficients (VRS input-orientation), the difference is 0.0704 (0.7315 – 0.6610) in the box industry. By comparison, in the gold industry, it is 0.1108 (0.8503 – 0.7395). A similar pattern is also found in the MAD, which varies more in the gold industry between the *DP* and *NPPE* compared to *GPPE* than the box industry. For example, in Table 5 - 14, input-orientation, for the box industry, the difference between the MAD is 0.0019 (0.0137 – 0.0118). In contrast, the difference is 0.0349 (0.1050 – 0.0701) in the gold industry. The

[36] This information is gained from manually checking the annual reports of the sample firms. For instance, in the FY 2019 annual report of AngloGold Ashanti LTD, page 36, the report stated "the majority of mining assets are amortised using the units-of-production method where the mine operating plan calls for production from a well-defined proved and probable Ore Reserve" (Anglo Gold Ashanti, 2019, p. 36).

change in discriminatory power is also evident in Table 5 - 13. The mean (median) of the box industry is approximately 0.97 (0.97) under the VRS input-orientation model. However, for the gold industry, the mean (median) is approximately 0.77 (0.81). The box plot in Figure 5 - 4 also illustrates this finding graphically. The plots for box industries on the left are highly similar when using alternative accounting variables. By comparison, the difference between the three alternative accounting variables is more obvious in the gold industry.

In summary, this section (section 5.3.2) tests whether the simulation's findings are consistent with the result using archival data. The results show that the relationships in the simulation hold when the archival data are generally homogeneous and consistent with the simulated assumptions of zero residual value, similar age of assets, and the straight-line method. In relatively more heterogeneous samples, greater differences depend on the choice of accounting measure reflecting specific industry features and accounting method choices.

5.3.3. Summary

This test examined how alternative forms of accounting variables, i.e., stock and flow form accounting variables, impact the Financial DEA results. The tests demonstrate how stock and flow forms of accounting variables (*GPPE, NPPE, DP*) are impacted by accounting choices and operational characteristics. As a result, the Financial DEA results using *NPPE* and *DP* diverged from the comparison point, the Financial DEA results using *GPPE*. The degree of divergence is influenced by the extent of variation of the factors of interest and sample sizes. When the factors have large variations, the Financial DEA results using *NPPE,* and *DP* diverge further away from the comparison point. However, with relatively large sample sizes, the differences between Financial DEA results using *NPPE* or *DP* compared to *GPPE* are lessened.

However, in the empirical test with archival data, the findings in the simulated test hold depending on the specific industry features. The box industry is relatively homogeneous and has similar features to the simulated sample. The *GPPE* generated the most similar efficiency measures, followed by *DP* and *NPPE*. The gold industry is relatively heterogeneous and has very different features from the simulated sample. The *GPPE* is relatively heterogeneous due to the uniqueness of the ore sites and machinery requirements. *DP* is heterogeneous since this industry commonly chooses the units-of-activity depreciation method, and the depreciation amount charged each year can vary significantly. *NPPE* is, by comparison, more similar to

the *GPPE* since most of the plant is relatively new, and the portion of the non-depreciable plant is relatively high. Therefore, the Financial DEA results using *NPPE* are closer than those using *DP* compared to those using *GPPE*.

5.4. Test Three: Alternative Accounting Variables

This test examines the impact of different degrees of disturbances due to alternative accounting variables on the construct validity of Financial DEA results. In this test, first, variables are altered within the same set of indicators. Second, variables are altered based on the alternation of indicators. As discussed in section 4.4.3, this test investigates, under the formative modelling lens, the impact of alternative accounting variables on construct validity when the indicators are fixed (convergent validity) or when the indicators vary (discriminant validity). As introduced by section 3.3.3, convergent validity is defined as the convergence of measures intended to represent the same construct; discriminant validity is defined as the divergence among measures designed to represent different constructs (Campbell, 1960; Campbell & Fiske, 1959; Edwards, 2003).

This section develops the propositions to test the potential effect of alternative accounting variables on the convergent and discriminant validity of Financial DEA results.

> ***P*$_{3a}$.** *For the same set of indicators, alternative accounting variables generate the same constructs of firm performance, showing convergent validity.*
>
> ***P*$_{3b}$.** *For different sets of indicators, alternative accounting variables generate different constructs of firm performance, showing discriminant validity.*

Under formative modelling, as discussed in section 2.2.3, Financial DEA is used to benchmark DMUs as a Multiple-Criteria Decision Making (MCDM) tool (Cook et al., 2014; da Silva et al., 2021; Martín-Gamboa et al., 2017). The inputs and outputs are two sets of performance criteria. The inputs are to be minimised, representing the less-the-better types of measures for benchmarking, instead of factors of production used in productive efficiency models. The outputs are to be maximised, representing the more-the-better measures for benchmarking, instead of the output volume used in productive efficiency models. Unlike reflective modelling, which forms a production frontier of an underlying production process,

formative modelling forms a best-practice frontier without an underlying production process (Cook et al., 2014).

5.4.1. Empirical Analysis

5.4.1.1. Empirical data

The data used in this test were collected for the reasons discussed in section 3.3.3.2. As reported in Table 5 - 15 Panel A, the initial samples were 388 (automobile), 59 (box), 217 (clothing), 360 (food), 409 (gold), and 55 (personal services). Next, observations with missing data were excluded. The firms with zero values were also excluded due to the algorithm features of DEA (Charnes et al., 1986; Golany & Roll, 1989). The final sample sizes were 253 (automobile), 53 (box), 160 (clothing), 218 (food), 43 (gold), and 47 (personal services).[37]

Table 5 - 15 Test Three Archival Sample Description

Panel A: Data Cleaning Process						
Industry	**Auto-mobile**	**Box**	**Clothing**	**Food**	**Gold**	**Personal Services**
All COMPUSTAT firm-years (2015-2019)	388	59	217	360	409	55
Delete observations missing values	-97	-6	-42	-81	-91	-5
Delete observations with zero values	-38	-0	-15	-61	-275	-3
Full sample	**253**	**53**	**160**	**218**	**43**	**47**
Panel B: Negative Value Translation						
Adjustment on CEQ	+5635	+1541	+319	+687	+0	+1407
Adjustment on NI	+4019	+421	+337	+10396	+3094	+79

Note: (a) Detailed composition of each industry can be found in Appendix 8. (b) Values in Panel B are in the unit of millions United States dollars (USD).

It is worth noting that the gold industry has a large proportion of zero values (275), which reduced the sample size from 396 to 43. As noted in the industry selection process, the gold industry was chosen to represent a relatively heterogeneous industry. However, due to the differences in variables, database, and time length, the final sample of the gold industry

[37] The box and gold industries have different final sample size than Test Two (section 5.3.2.1) due to different accounting variable required in the data cleaning process. Also, different from the test in section 5.3, this test did not bootstrap the data since the focus of this test is to compare the impact of various models on the Financial DEA results within various industries. The dimensionality across models was not constant. Although bootstrapped data would provide each industry the same number of DMUs, this test does not focus on the comparison across industries using one model. Rather, this test focuses on the change of models within one industry. Various dimensionality of models would lead to incomparable discriminatory power models and bootstrapping does not contribute the research purpose.

differs from Demerjian et al. (2018). Given that such a high portion of data was deleted, the sample set consists of several firms over the five financial years with more homogeneous features than in Demerjian et al. (2018).

Next, all variable values were inflated to the financial year of 2019 using the consumer price index (CPI) (Rouse & Tripe, 2016). Last, any negative values (regarding Common and Ordinary Equity – *CEQ* and Net Income – *NI*) were translated to positive values by adding a positive constant value to each DMU so that any negative values were translated to be positive in the DEA algorithm (Bowlin, 1999; Charnes et al., 1983; Seiford & Zhu, 2002) (Table 5 - 17 Panel B).

The final sample meet the expectation of various sample sizes discussed in section 3.3.3.2. As reported in Table 5 - 15, relatively large industries such as the automobile (N = 253) and food industries (N = 218) provide a comparison with relatively small industries such as box (N = 53), gold (N = 43) and personal services (N = 47).

Also, the descriptive statistics of the final sample from the six industries meet the expectation of various industrial features discussed in section 3.3.3.2. For instance, the clothing industry is relatively labour intensive because of sales personnel. In Table 5 - 18, the clothing industry has the highest ratio of operating expenses (*XOPR*) to total assets (*AT*) (1.01), which indicates that for every dollar of total resources used by a firm, the operating expense is 1.01 dollars. By comparison, the food and gold industries are relatively capital intensive since they rely on production machinery. In Table 5 - 18, the food and gold industries have a ratio of the net property plant and equipment (*NPPE*) to total assets (*AT*) of 0.47 and 0.65, respectively. The ratios indicate that the cost of capital takes approximately 47% and 65% of the total resources in the two industries.

Moreover, industries were selected that had varying firm sizes. For example, the DMUs from the automobile industry tend to be relatively large, considering the machinery and scope requirements. In Table 5 - 18, the mean of total assets (*AT*) is approximately 16,000 million USD. In contrast, DMUs from the personal services industry are more likely to be small and medium-size businesses due to their business nature, such as personal laundry and beauty shops. In Table 5 - 18, the mean of the total assets (*AT*) is approximately 3,000 million USD, which is approximately 20% of the automobile industry.

5.4.1.2. ***Empirical models***

In Test A, to test P_{3a}, four Financial DEA models were developed based on the previous studies. In formative modelling, the same set of indicators defines the same construct. In Test A, the indicators were assets and expenses on the input side and sales on the output side. According to the convergent validity of constructs, all constructs measured by the different models should exhibit convergence through disturbances, and measurement errors arise between the measures and the indicators (Edwards, 2003). Therefore, the different models in Test A were expected to show convergent validity.

Model I was adapted from Demerjian et al. (2012) as a comparison point for the other models. The Financial DEA models used by Demerjian et al. (2012) had seven inputs: the cost of goods sold (*COGS*), Selling, General and Administrative Expense (*XSGA*), Net Property, Plant and Equipment (*NPPE*), capitalised operating lease expenses (*MRC1-5* discounted at the rate of 10%), five-year capitalised Research and Development Expense (*XRD*), Goodwill (*GDWL*), and other intangibles (intangible assets, *INTAN* less *GDWL*). On the output side, sales (*SALE*) was the sole output. Model I adapted this model to reflect data availability and provide a better comparison point for the other models. Accordingly, it excluded the operating lease and the research and development expenses due to the high volume of missing data. Unlike Demerjian et al. (2012) and Demerjian et al. (2013), goodwill was not separated from intangible assets because of data availability. Model I also included current assets (*ACT*) as an input because including *ACT* enabled the aggregated model (Model IV).

Model II examined whether partially covered indicators influence the convergent validity of the Financial DEA results. Compared to Model I, Model II replaced the *XSGA* (the selling, general and administrative expense) with *XOPR* (the operating expense). In accounting reports, *XSGA* is generally a sub-category of *XOPR* and only partially covers the information in *XOPR*.

Model III examined whether the omission of variables impacts the convergent validity of the Financial DEA results. Compared to Model I, Model III omitted the intangible assets (*INTAN*), a sub-indicator of the assets. If *INTAN* covers unique information, omitting *INTAN* from the total assets should weaken the convergent validity of Financial DEA results.

Table 5 - 16 Test Three A Model Specification

	Indicators	Model I	Model II	Model III	Model IV
		Variables			
Inputs	**Assets**	*NPPE, INTAN, ACT*	*NPPE, INTAN, ACT*	*NPPE, ACT*	***AT***
	Expenses	*XSGA, COGS*	***XOPR***, *COGS*	*XSGA, COGS*	***XSGA + COGS***
Outputs	**Sales**	*SALE*	*SALE*	*SALE*	*SALE*

Note: (a) *AT*, total assets. *ACT*, current assets. *NPPE*, net Property, Plant and Equipment. *INTAN*, intangibles. *XOPR*, operating expenses. *XSGA*, selling, general and administrative expenses. *COGS*, cost of goods sold. *LT*, total liabilities. *NI*, net income. *CEQ*, common equity. *MKVALT*, market value. *SALE*, sales.

Table 5 - 17 Test Three B Model Specification

	Indicators	Variables	Model I	Model II	Model III	Model IV	Model V	Model VI	Model VII	Model VIII
Inputs	**Assets**	*AT*	√	√	√	√	√	√		
	Liability	*LT*							√	
	Expenses	*XOPR, COGS*	√	√	√	√	√	√		
	Equity	*CEQ*				√	√	√	√	
	Sales	*SALES*								√
	Profit	*NI*								√
Outputs	**Sales**	*SALES*	√		√	√		√		
	Profit	*NI*		√	√		√	√	√	
	Equity	*MKVALT*								√

Note: (a) Model I was adapted from Demerjian et al. (2012) and Demerjian et al. (2013). (b) Model II was the transit model between Model I and Model III. (c) Model III was adapted from Harrison and Rouse (2016). (d) Model IV was the DuPont ratio model adapted from Feroz et al. (2001,2003). (e) Model V was the transit model between Model IV and VI. (f) Model VI was adapted from the stage one DEA model from Seiford and Zhu (1999) and Zhu (2000). (g) Model VII was the funding model adapted from Smith (1990). (h) Model VIII was adapted from the stage two DEA model from Seiford and Zhu (1999) and Zhu (2000).

Model IV aimed to test the impact of aggregating accounting variables on the convergent validity of Financial DEA. Compared to Model I, in Model IV, the *NPPE, INTAN, and ACT* were aggregated into *AT* (total assets) or the assets indicator. To represent the indicator of expenses, the *XSGA* and *COGS* were aggregated into one variable.

In Test B, to test P_{3b}, eight Financial DEA models were designed (Table 5 - 17). In Test B, to operationalise one indicator, the type of accounting variable is fixed. The only reason for accounting variables to alter is because of a change in indicators. Different sets of indicators define different constructs, according to formative modelling. The constructs formed and measured by different models in Test B were expected to show discriminant validity (Edwards, 2003).

In Test B, Model I was adapted from Demerjian et al. (2012) and Demerjian et al. (2013) for financial efficiency measurement. This model served as a comparison point for the other models in this test. As discussed in Test A, *XOPR* and *COGS* were used as accounting variables measuring the expenses. Unlike Test A Model I, *XOPR* was used here instead of *XSGA* to ensure comparability with the other models in Test B. Unlike Test A Model I, this model merged the *NPPE, INTAN, and ACT* into *AT* for comparability with the other models in Test B.

Model II was based on Model I. However, Model II substitutes *NI* for *SALES* as the only output. Therefore, Model II forms a different construct from Model I.

Model III was adapted from Harrison and Rouse (2016) to measure financial efficiency. Harrison and Rouse (2016) suggested the inputs should be *AT, XOPR,* and the outputs should be *SALES* and *NI*. Additionally, this test included *COGS* as the third input for comparability with the other models. This model formed a different construct from Model I by including the profit indicator as an attribute for output quality.

Model IV was adapted from Feroz et al. (2003) to measure financial efficiency based on the DuPont accounting ratio. Feroz et al. (2003) used *TA, CEQ* (common equity), and costs as inputs and the *SALES* as the output. This test modified Feroz et al. (2003) using *XOPR* and *COGS* as inputs for comparability with the other models. Compared to Model I, Model III included equity as an attribute input contributing to financial performance.

Model V was based on Model IV. However, Model V substitutes *NI* for *SALES* as the only output. Therefore, Model V forms a different construct from Model I.

Model VI was adapted from the first stage DEA model in Seiford and Zhu (1999) and Zhu (2000). In these studies, the DEA model incorporated the number of employees, assets, and stockholder's equity as inputs to generate revenue and profits. To make the model comparable to the other models in Test B, this test adapted the inputs as assets (*AT*), expenses (*XOPR*, *COGS*) and equity (*CEQ*) to generate outputs, sales (*SALES*), and profit (*NI*).

Model VII was adapted from Smith (1990), aiming to measure a firm's efficiency to generate profits from different funding sources. Smith (1990) used *CEQ* and *LT* (total liabilities) to generate earnings before tax and interest (*EBIT*), interest expenses and tax expenses. This test modified the outputs to *NI* as it is essentially the *EBIT* minus interest expenses and tax expenses.

Model VIII was adapted from the second stage DEA model in Seiford and Zhu (1999) and Zhu (2000). The DEA model used revenue and profits to generate *MKVALT* (market value), total return index, and earnings per share in the studies. This test modified the outputs to *MKVALT* only since the other two measures (i.e. return to investors and earnings per share) are not accounting variables.

In sum, in Test B, Models I to VII measure slightly different constructs relating to financial efficiency. The constructs in Models I to VII were expected to diverge, given their similarity. In contrast, Model VIII measured a distinct construct due to the selection of indicators. Therefore, it was expected that Model VIII's level of discriminant validity would be higher than Models I to VII.

All the models specified above were analysed using variable return to scale (VRS) models since the empirical data includes DMUs with a wide scale. For example, in Table 5 - 18, the total assets (*AT*) in the automobile industry ranges from approximately 1.43 million USD to 268,120.32 million USD. The constant return to scale (CRS) was not a suitable choice for this test since the CRS model does not incorporate economies of scale effects (Banker et al., 1984; Charnes et al., 1978).

5.4.1.3. Empirical Results

Descriptive statistics – variables

Table 5 - 18 reports the descriptive statistics of the archival data. In general, the results indicate that the selected industries provide a range of research settings suitable for the research purpose. First, industries exhibit various levels of homogeneity. For the selected accounting variables, based on the coefficient of variation, the box industry is the most homogeneous as all coefficients of variation are below one. By comparison, other industries are relatively heterogeneous. The personal services industry, gold industry, food industry, and clothing industry are relatively more heterogeneous, increasing the coefficient of variation. Moreover, the automobile industry is the most heterogeneous since the coefficient of variation is often approximately three. As discussed in section 5.4.1.1, the samples were selected to represent a wide variety of industry heterogeneity.

Second, the sample industries are of different sizes (N), so the impact of sample size on the Financial DEA results can be assessed. For example, the largest industry has 253 DMUs (automobile industry), which is about six times the smallest industry's size with 43 DMUs (gold industry).

Table 5 - 18 Test Three Descriptive Statistics of Archival Data

Panel A: Automobile Industry N = 253									
	MIN	**Q1**	**MED**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**	**/AT**
AT	1.43	435.83	1847.20	5302.86	268120.32	16102.50	50183.15	3.12	1.00
ACT	0.83	275.18	845.67	2696.66	120538.08	6671.32	20781.87	3.12	0.41
INTAN	0.00	47.33	237.56	719.10	31172.18	1319.60	4305.63	3.26	0.08
NPPE	0.05	81.63	446.76	1315.46	83963.34	4847.40	15016.18	3.10	0.30
COGS	1.28	563.53	1694.24	5759.42	140881.42	10665.24	28645.66	2.69	0.66
XOPR	2.69	673.37	2098.61	6451.64	157301.01	11903.92	31728.23	2.67	0.74
XSGA	1.04	68.87	249.00	715.00	24593.48	1238.68	3230.19	2.61	0.08
LT	0.51	250.83	999.27	4197.43	231703.68	12752.65	41218.00	3.23	0.79
CEQ-PO	0.79	5782.57	6155.99	7209.37	69031.46	8854.63	9237.49	1.04	0.55
MKVALT	0.93	323.31	1458.07	5345.86	97546.13	6596.15	13645.37	2.07	0.41
SALE	1.53	726.37	2353.52	7415.81	176913.78	13473.66	35879.96	2.66	0.84
NI-PO	0.44	4022.77	4099.75	4353.63	14577.83	4523.46	1616.48	0.36	0.28
Panel B: Box Industry N = 53									
	MIN	**Q1**	**MED**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**	**/AT**
AT	706.69	4639.15	6701.09	10921.80	30156.70	9177.51	7304.47	0.80	1.00
ACT	322.71	1216.70	1805.60	3196.96	4974.30	2255.84	1345.92	0.60	0.25
INTAN	50.31	1258.90	1829.88	3673.02	11345.10	3116.06	2843.46	0.91	0.34
NPPE	270.26	1258.50	2989.20	3456.60	11189.50	3113.84	2583.61	0.83	0.34
COGS	591.05	3395.82	4692.31	7358.28	13432.90	5392.90	3062.87	0.57	0.59

XOPR	711.38	3772.86	5167.32	7771.38	15148.10	5916.15	3338.88	0.56	0.64
XSGA	120.33	378.99	491.92	554.96	1715.20	523.25	337.89	0.65	0.06
LT	676.36	2867.41	4214.30	10447.65	18470.60	6953.00	4976.70	0.72	0.76
CEQ-PO	0.80	2038.58	2670.96	3343.68	14241.46	3670.86	3228.41	0.88	0.40
MKVALT	32.42	2930.53	4832.61	8470.62	20994.15	6237.72	4771.88	0.77	0.68
SALE	785.60	4534.62	6258.45	9278.94	18289.00	7027.37	4007.35	0.57	0.77
NI-PO	0.92	583.55	683.14	884.08	2365.22	743.77	378.08	0.51	0.08

Panel C: Clothing Industry N = 160

	MIN	Q1	MED	Q3	MAX	MEAN	SD	CV	/AT
AT	2.07	291.84	1703.90	3223.12	31342.00	3299.68	5113.34	1.55	1.00
ACT	1.23	210.06	941.04	1800.60	20556.00	1704.60	3121.69	1.83	0.52
INTAN	0.01	38.08	229.98	683.06	8014.98	816.98	1556.27	1.90	0.25
NPPE	0.24	36.23	236.15	767.55	7963.00	579.62	1004.21	1.73	0.18
COGS	0.60	251.94	1170.93	1963.61	21775.52	1972.37	3580.12	1.82	0.60
XOPR	6.81	380.76	2017.39	3566.99	34985.60	3341.24	5843.22	1.75	1.01
XSGA	3.88	117.15	708.72	1485.70	13210.08	1368.87	2314.38	1.69	0.41
LT	0.87	179.52	528.60	1767.81	23287.00	1777.97	3123.21	1.76	0.54
CEQ-PO	0.37	451.33	4.00	2153.63	13842.63	1835.01	2227.55	1.21	0.56
MKVALT	2.40	165.93	2066.67	6625.40	153587.64	8018.49	21404.78	2.67	2.43
SALE	0.69	401.92	2291.45	4422.42	40681.68	3896.86	6817.40	1.75	1.18
NI-PO	0.21	342.55	428.09	660.32	4958.60	621.45	663.38	1.07	0.19

Panel D: Food Industry N = 218

	MIN	Q1	MED	Q3	MAX	MEAN	SD	CV	/AT
AT	1.30	382.88	3091.31	10284.13	134040.57	10781.57	21368.50	1.98	1.00
ACT	0.94	125.31	916.59	2175.57	23793.61	2190.12	3931.24	1.79	0.20
INTAN	0.00	17.88	1235.70	4014.88	114636.39	6305.20	16791.97	2.66	0.58
NPPE	0.97	802.87	1723.59	3604.23	69362.36	5043.52	10160.72	2.01	0.47
COGS	0.51	153.78	1968.02	6875.26	66680.36	6305.35	11228.13	1.78	0.58
XOPR	3.44	405.76	2614.51	7135.45	70651.62	7049.58	12734.62	1.81	0.65
XSGA	1.96	63.60	331.72	1474.47	7869.80	914.72	1280.14	1.40	0.08
LT	0.16	10397.07	10480.85	11016.31	21834.96	10832.48	1373.68	0.13	1.00
CEQ-PO	1.18	307.31	2098.80	5517.18	68460.72	6134.86	12091.16	1.97	0.57
MKVALT	5.52	297.96	2978.30	13809.65	112596.05	10489.23	17747.24	1.69	0.97
SALE	1.42	460.89	2957.96	8286.34	73795.18	8076.32	13699.28	1.70	0.75
NI-PO	0.01	66.21	691.81	2187.96	11077.00	1704.83	2471.87	1.45	0.16

Panel E: Gold Industry N = 43

	MIN	Q1	MED	Q3	MAX	MEAN	SD	CV	/AT
AT	94.81	1809.75	7117.20	8932.94	44392.00	9050.87	10718.61	1.18	1.00
ACT	48.78	481.39	767.02	1727.28	6887.00	1573.74	1892.78	1.20	0.17
INTAN	0.02	15.17	158.80	488.67	4783.00	466.58	871.40	1.87	0.05
NPPE	17.18	1008.88	5175.06	6917.00	25276.00	5903.46	6174.42	1.05	0.65
COGS	2.09	259.21	908.07	2045.62	6348.16	1524.07	1632.99	1.07	0.17
XOPR	4.39	316.35	1127.65	2256.95	6976.00	1717.57	1817.37	1.06	0.19
XSGA	2.30	37.59	111.00	256.96	728.00	193.51	202.80	1.05	0.02
LT	12.88	527.28	2770.95	4106.54	18369.77	4056.70	5278.20	1.30	0.45
CEQ-PO	70.21	1273.96	4185.20	5078.72	21432.00	4477.97	4697.11	1.05	0.49
MKVALT	98.00	1153.64	2480.81	9516.87	35107.60	6242.54	8105.87	1.30	0.69
SALE	1.13	494.32	1612.20	3301.88	9848.15	2585.52	3059.37	1.18	0.29
NI-PO	0.58	2937.79	3100.51	3241.09	7063.00	3104.24	1085.17	0.35	0.34

Panel F: Personal Services Industry N = 47									
	MIN	**Q1**	**MED**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**	**/AT**
AT	2.03	929.76	1295.84	3316.27	13677.43	3165.31	3996.32	1.26	1.00
ACT	0.65	164.75	295.39	483.69	3086.74	608.48	791.49	1.30	0.19
INTAN	0.10	429.55	701.62	1046.17	3605.37	1021.22	982.65	0.96	0.32
NPPE	0.39	52.90	249.22	645.96	2123.53	529.70	657.77	1.24	0.17
COGS	1.08	535.03	972.26	1663.85	3636.97	1227.19	1023.77	0.83	0.39
XOPR	9.58	687.40	1152.81	2336.11	5652.80	1628.28	1465.27	0.90	0.51
XSGA	8.19	84.76	180.55	457.74	2015.83	401.09	539.50	1.35	0.13
LT	4.51	544.61	903.19	2935.89	11913.27	2604.22	3442.29	1.32	0.82
CEQ-PO	0.73	1455.11	1781.10	2105.69	4642.20	1966.82	1015.07	0.52	0.62
MKVALT	11.67	543.46	1458.17	5498.19	25642.78	3944.06	5936.69	1.51	1.25
SALE	2.60	795.41	1497.77	3215.75	7085.12	2013.11	1866.51	0.93	0.64
NI-PO	0.62	81.69	133.26	391.20	981.68	263.43	273.69	1.04	0.08

Note: (a) *AT*, total assets. *ACT*, current assets. *NPPE*, net Property, Plant and Equipment. *INTAN*, intangibles. *XOPR*, operating expenses. *XSGA*, selling, general and administrative expenses. *COGS*, cost of goods sold. *LT*, total liabilities. *NI-PO*, net income, translated to positive values. *CEQ-PO*, common equity, translated to positive values. *MKVALT*, market value. *SALE*, sales. (b) The descriptive features include minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the maximum value (MAX), the mean value (MEAN), the standard deviation (SD), the coefficient of variation, which is the ratio of the standard deviation to the mean (CV), and the proportion to total assets (/AT). (c) Values are in the unit of millions of United Sates dollars (USD).

Third, the industries have different characteristics. For instance, the gold industry has a relatively higher rate of property plant and equipment cost as a proportion of total assets (*NPPE* to *AT*), which is approximately 0.65. This is followed by the food industry, which is approximately 0.47 compared to personal services, approximately 0.17. The food industry also has a relatively higher percentage of intangible assets as a proportion of total assets (*INTAN* to *AT*), approximately 0.58. For expenses, the clothing industry has a relatively higher proportion of operation expenses (*XOPR*) to *AT*, approximately 1. For the clothing industry, the general expenses also have a relatively higher proportion, approximately 0.41 of the *AT*. For liabilities, the food industry has the highest leverage rate at approximately 1, which is the ratio of total liabilities (*LT*) to *AT*, followed by the personal services industry (0.82) and the automobile industry (0.79).

Efficiency scores – Test A

Table 5 - 19 reports the efficiency scores of Test A. The results of the VRS output orientation are reported below. The VRS input orientation produces similar results (untabulated). The distribution of efficiency scores is also depicted in Figure 5 - 5.

Table 5 - 19 Test Three A Efficiency Score

Panel A: Automobile Industry N = 253										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9502	40	98	0.3456	0.6544	0.9275	0.9745	1.0000	0.0669	0.0704
Model II	0.9442	26	87	0.3456	0.6544	0.9206	0.9643	1.0000	0.0682	0.0722
Model III	0.9042	31	63	0.4881	0.5119	0.8534	0.9402	0.9880	0.1065	0.1178
Model IV	0.8945	8	29	0.5850	0.4150	0.8770	0.9128	0.9596	0.0985	0.1101
Panel B: Box Industry N = 53										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9874	14	35	0.0691	0.9309	0.9750	1.0000	1.0000	0.0193	0.0195
Model II	0.9843	14	32	0.0691	0.9309	0.9727	0.9958	1.0000	0.0210	0.0213
Model III	0.9870	12	34	0.0691	0.9309	0.9750	0.9989	1.0000	0.0192	0.0195
Model IV	0.9663	5	16	0.0942	0.9058	0.9437	0.9713	0.9923	0.0290	0.0300
Panel C: Clothing Industry N = 160										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9497	30	70	0.3837	0.6163	0.9205	0.9719	1.0000	0.0656	0.0691
Model II	0.9263	23	47	0.5487	0.4513	0.8808	0.9436	0.9994	0.0789	0.0852
Model III	0.9349	21	57	0.4758	0.5242	0.9036	0.9562	1.0000	0.0806	0.0862
Model IV	0.8455	9	13	0.7599	0.2401	0.7969	0.8381	0.9050	0.1026	0.1214
Panel D: Food Industry N = 218										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9448	41	97	0.5127	0.4873	0.9244	0.9794	1.0000	0.0886	0.0938
Model II	0.9399	43	91	0.5127	0.4873	0.9140	0.9734	1.0000	0.0906	0.0964
Model III	0.9071	21	48	0.5451	0.4549	0.8647	0.9310	0.9862	0.1044	0.1151
Model IV	0.8777	14	34	0.9236	0.0764	0.8374	0.9037	0.9668	0.1330	0.1515
Panel E: Gold Industry N = 43										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9224	23	25	0.6480	0.3520	0.8997	1.0000	1.0000	0.1517	0.1645
Model II	0.9206	21	25	0.6480	0.3520	0.8997	0.9960	1.0000	0.1516	0.1646
Model III	0.9142	22	24	0.6759	0.3241	0.8997	1.0000	1.0000	0.1637	0.1791
Model IV	0.8202	9	11	0.8430	0.1570	0.7432	0.8852	0.9875	0.2119	0.2583
Panel F: Personal Services Industry N = 47										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9811	32	34	0.1776	0.8224	0.9837	1.0000	1.0000	0.0405	0.0413
Model II	0.9764	27	31	0.1776	0.8224	0.9736	1.0000	1.0000	0.0435	0.0445
Model III	0.9449	25	27	0.4833	0.5167	0.9015	1.0000	1.0000	0.0924	0.0977
Model IV	0.8982	10	12	0.7057	0.2943	0.8735	0.9458	0.9930	0.1489	0.1658

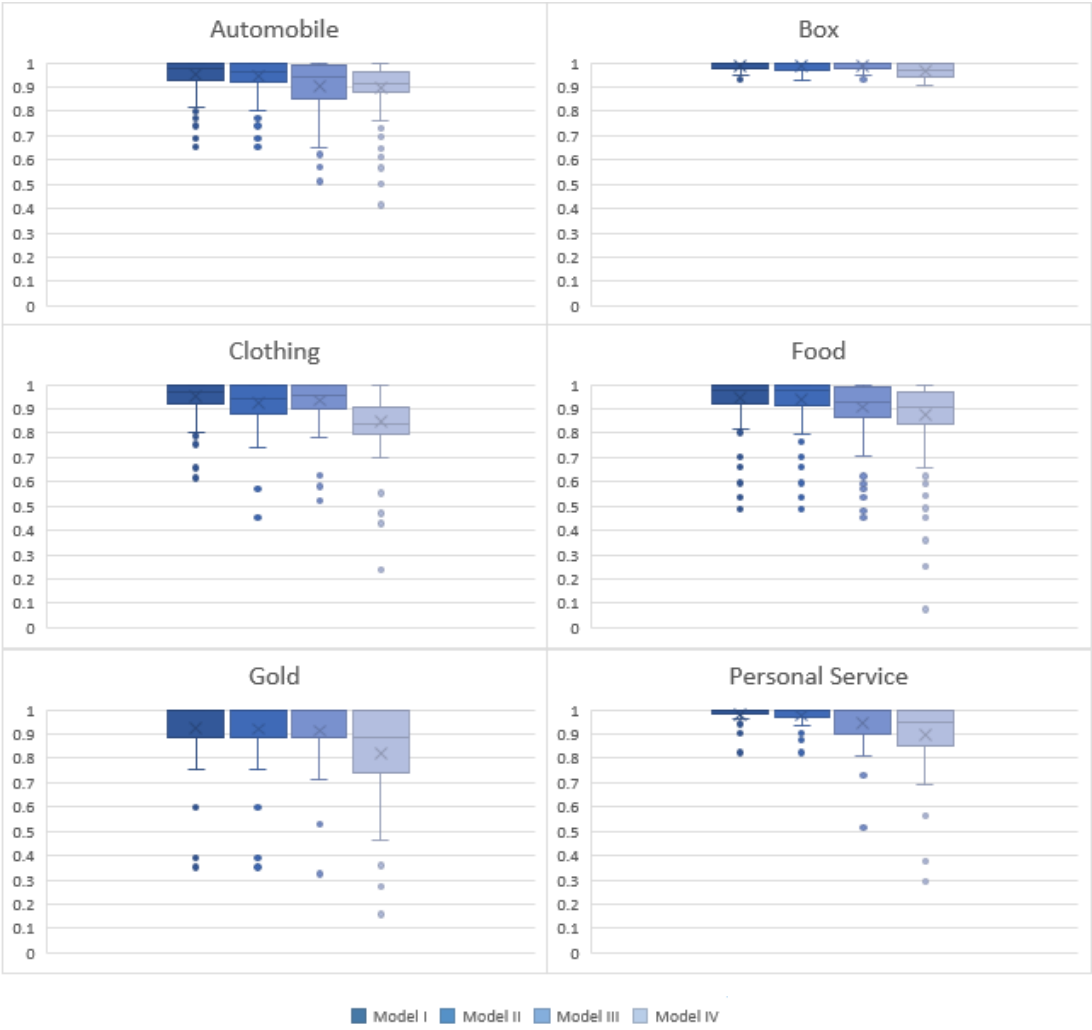
Note: (a) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs that are at least 99% efficient (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (b) DEA models were the variable return of scale output-orientation. The variable return of scale, input-orientation models, generated similar results, untabulated. (3) Model I was adapted from Demerjian et al. (2012) and Demerjian et al. (2013). Model II was designed to examine the effect of partially covered indicators. Model III excluded the intangible assets. Model IV used the aggregated variables.

Table 5 - 19 shows that the mean efficiency score is above 0.8 for all industries, reflecting the sample selection process, which focused on homogeneity. The high mean efficiency scores indicate that the selected industries are relatively homogeneous.[38]

The results (Table 5 - 19 and Figure 5 - 5) suggest that the homogeneity level of the industry impacts the level of convergent validity. Relatively homogeneous industries are expected to have a relatively narrow range of efficiency scores, relatively low coefficient of variation and relatively more fully efficient DMUs. By contrast, relatively heterogeneous industries are expected to have a relatively broad range of efficiency scores, a relatively high coefficient of variation, and relatively fewer fully efficient DMUs. As illustrated in Table 5 - 19 and Figure 5 - 5, the most homogeneous industry is the box industry. The food, automobile and clothing industries are, by comparison, less homogeneous.

As illustrated in Figure 5 - 5, the efficiency scores of the four models in Test A are similar, suggesting convergent validity consistent with the process of model construction. However, the efficiency scores generated by Model IV are the least similar to the others, demonstrating the effect of aggregating accounting variables

[38] The average efficiency score of the gold industry is quite different compared to Test Two (section 5.3.2.3), where the average efficiency score of the gold industry is around 0.8. In this test, the average efficiency score is around 0.9. The difference of average efficiency scores is due to different variables required by the models, by Test Three and Test Two.



.

Figure 5 - 5 Test Three A Box Plot for the Distribution of Efficiency Scores

Variation in Financial DEA results – Test A

Table 5 - 20 reports Test A results and summarises the correlations among all the models. Test A aims to examine P_{3a} for the same set of indicators and whether alternative accounting variables generate the same Financial DEA construct and demonstrate convergent validity. Since the indicators are the same, the relative domain of the construct should be the same. However, disturbance on the construct can result from measurement errors on individual measures that diverge from the domain of interest (Mackenzie et al., 2005).

Table 5 - 20 Test Three A Variation in Financial DEA Results

Spearman's Ranking Correlations

Panel A: Automobile Industry N = 253

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.9455	1.0000		
Model III	0.7848	0.7310	1.0000	
Model IV	0.6877	0.7263	0.7636	1.0000

Panel B: Box Industry N = 53

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.8481	1.0000		
Model III	0.9435	0.7897	1.0000	
Model IV	0.6908	0.8103	0.6953	1.0000

Panel C: Clothing Industry N = 160

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.7544	1.0000		
Model III	0.8552	0.6594	1.0000	
Model IV	0.6092	0.6998	0.6014	1.0000

Panel D: Food Industry N = 218

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.9710	1.0000		
Model III	0.7825	0.7839	1.0000	
Model IV	0.6622	0.6660	0.6016	1.0000

Panel E: Gold Industry N = 43

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.9739	1.0000		
Model III	0.9738	0.9461	1.0000	

Pearson Correlations

Panel A: Automobile Industry N = 253

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.9755	1.0000		
Model III	0.8023	0.7632	1.0000	
Model IV	0.7501	0.7531	0.8412	1.0000

Panel B: Box Industry N = 53

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.9597	1.0000		
Model III	0.9951	0.9524	1.0000	
Model IV	0.7157	0.7888	0.7119	1.0000

Panel C: Clothing Industry N = 160

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.8340	1.0000		
Model III	0.9132	0.7729	1.0000	
Model IV	0.6963	0.7378	0.7553	1.0000

Panel D: Food Industry N = 218

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.9929	1.0000		
Model III	0.9155	0.9186	1.0000	
Model IV	0.6950	0.6911	0.6724	1.0000

Panel E: Gold Industry N = 43

	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.9985	1.0000		
Model III	0.9955	0.9942	1.0000	

Model IV	0.7603	0.7735	0.7671	1.0000
Panel F: Personal Services Industry N = 47				
	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.8589	1.0000		
Model III	0.6923	0.5199	1.0000	
Model IV	0.5044	0.5477	0.6831	1.0000

Model IV	0.8626	0.8646	0.8677	1.0000
Panel F: Personal Services Industry N = 47				
	Model I	*Model II*	*Model III*	*Model IV*
Model I	1.0000			
Model II	0.8889	1.0000		
Model III	0.4325	**0.3408**	1.0000	
Model IV	**0.1776**	**0.2123**	0.8195	1.0000

Note: (a) The numbers in bold have p-values greater than 0.01. (b) The numbers in grey shadows are greater than 0.8, which demonstrate convergent validity. (c) DEA models were the variable return of scale, output-orientation. The variable return of scale, input-orientation models, generated similar results, untabulated. (d) Model I was adapted from Demerjian et al. (2012) and Demerjian et al. (2013). Model II was designed to examine the effect of partially covered indicators. Model III excluded the intangible assets. Model IV used the aggregated variables.

In Table 5 - 20, to demonstrate convergent validity, the correlations are expected to be highly, but not perfectly, correlated (correlation coefficient > 0.8, or significant p-values < 0.1). The numbers in bold have p-values greater than 0.01. The numbers in grey shadows are greater than 0.8, which demonstrate convergent validity. High statistical correlations illustrate high levels of convergent validity among the same constructs. The imperfect statistical correlations (correlation coefficient < 1.0) suggest that the measurement errors have impacted the domain of the constructs. Table 5 - 20 shows Spearman's ranking correlations and Pearson correlations. The DEA results generated using VRS output orientation are reported here. The VRS input-orientation results (untabulated) generated similar results.

Overall, in Table 5 - 20, Model II and Model III are highly correlated to Model I. This result confirms for P_{3a} that when the indicators are the same, alternative accounting variables define the same construct of firm performance, showing convergent validity. However, the magnitude of convergent validity is impacted by industry characteristics. Moreover, Model IV does not highly correlate to Model I in most cases, which suggests significant dimensionality change affects the discriminatory power of the DEA models. The key findings are threefold.

First, when the indicators are the same, alternative accounting variables define the same construct, showing convergent validity. As evident in Table 5 - 20, most cells of Model II and Model III are highly correlated against Model I.

Second, the magnitude of convergent validity is affected by the industry characteristics. In Model II, the composite construct for operating expense-intensive industries (e.g. clothing industry) tends to affect operating expense variables. Compared to Model I, Model II replaces *XSGA* with *XOPR*. The *XSGA* normally consists of non-production costs, such as accounting expenses, legal expenses, advertising, and marketing expenses. By comparison, *XOPR* contains the daily operating costs, which include the production cost and *XSGA*. In Table 5 - 18, the clothing industry has the most intensive operating expenses (*XOPR*/*TA* = 1.01). In Table 5 - 20, replacing *XSGA* with *XOPR* leads to a significant change of the composite construct measured by Financial DEA for the clothing industry. The change can also be found in weightings allocated by DEA. The mean weighting for *XSGA* is 0.4137 in Model I in the clothing industry, and for *XOPR* is 0.6352 in Model II (untabulated). By comparison, in the other models, the weightings on *XSGA* and *XOPR* are similar. For instance, in the

automobile industry, the mean weighting for *XSGA* is 0.0017 in Model I, and for *XOPR* is 0.0012 in Model II (untabulated).

In Model III, the composite construct is affected by omitting the intangible variable for the intangible intensive industries (e.g. food and personal services industries).[39] Model III omits the intangible assets, and the results demonstrate less convergent validity. For instance, in Table 5 - 18, the food industry (*INTAN/TA* = 0.58) and the personal services industry (*INTAN/TA* = 0.32) had the most intensive intangible assets. In Table 5 - 20, the correlation between Model I and III shows relatively less convergent validity for these industries, as evidenced by the coefficient of 0.7825 for the food industry and 0.6923 for the personal services industry. The weightings of intangible assets are 0.0473 and 0.2414 in the food and personal industry, respectively. These are relatively high weightings compared to the other industries. For instance, the intangible assets received a weighting of 0.0000 in the box industry (untabulated).

Another reason for Model III and Model I receiving less convergent validity is the low correlation between variables (Dyson et al., 2001). The correlation between intangible assets with the other inputs can lead to a significant change in efficiency scores. For instance, in Appendix 9, the correlation coefficients between intangible assets and the other inputs are relatively low (< 0.8) in the automobile industry. In Table 5 - 20, the correlations between Model III and Model I show relatively weaker convergent validity in the automobile (coefficient = 0.7858) than the other industries. Also, the weighting assigned by DEA on the intangible assets is 0.0687 in the automobile industry. This weighting is higher than the other inputs in Model I, and the *XSGA* has a weighting of 0.0017, the *COGS* has a weighting of 0.0250, the *ACT* has a weighting of 0.0065, and *NPPE* has a weighting of 0.0324 (untabulated).

Third, regarding the impact of dimensionality and discriminatory power, in Table 5 - 20, Model IV consistently exhibits low correlations with Model I, with almost all coefficients less than 0.8. Model IV aggregates accounting variables to three from six (Model I), reduces dimensionality, and increases discriminatory power (Dyson et al., 2001). Table 5 - 20 and

[39] Model III (dimensionality = 5) has different dimensionality from Model I (dimensionality = 6). However, compared to the difference of dimensionality between Model IV (dimensionality = 3) and Model I (dimensionality = 6), Model III is not impacted by the change of dimensionality as significantly as Model IV, compared to Model I. In the interpretation of Model III, the omission of intangibles was interpreted as the main reason for the change of the Financial DEA results.

Figure 5 - 5 show that Model IV has the smallest number of efficient DMUs in the models of all industries. This finding provides empirical evidence that the aggregated accounting variables can lead to quite different Financial DEA results than disaggregated accounting variables (Banker et al., 2007).

In sum, the results in Table 5 - 20 suggest that alternative accounting variables generate the same construct of firm performance, showing convergent validity when the indicators are the same. However, the magnitude of convergent validity is impacted by the industry characteristics and the dimensionality of Financial DEA models. The homogeneity level of industry and the aggregation of accounting variables in Financial DEA may impact the results.

Efficiency scores – Test B

Table 5 - 21 reports the descriptive statistics of the efficiency scores of Test B. The results of VRS output orientation are reported below. The other models produce similar results (untabulated). The distribution of efficiency scores is depicted in Figure 5 - 6. Overall, most mean efficiency scores are above 0.7, indicating relatively high homogeneity, consistent with the sample selection process. However, a few results have mean efficiency scores below 0.7, which relate primarily to Model VIII and Model II since these two models were designed very differently.

Comparing models, Model II, V, VII and VIII have very different descriptive statistics and graphical distributions across most industries. For example, for the personal services industry, the mean efficiency scores of these four models are all below 0.9. By comparison, the other models (Model I, III, IV, VI) all have mean efficiency scores above 0.9. Graphically, Figure 5 - 6 shows a clear division between these four models and the other models.

Table 5 - 21 Test Three B Efficiency Score

Panel A: Automobile Industry N = 253										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9053	12	37	0.4942	0.5058	0.8805	0.9163	0.9687	0.0910	0.1006
Model II	0.8576	5	27	0.9314	0.0686	0.7901	0.8901	0.9655	0.1361	0.1587
Model III	0.9561	21	63	0.2333	0.7667	0.9310	0.9648	0.9897	0.0402	0.0421
Model IV	0.9097	23	48	0.4942	0.5058	0.8844	0.9198	0.9802	0.0920	0.1011
Model V	0.8658	7	29	0.9309	0.0691	0.7943	0.9068	0.9663	0.1305	0.1508
Model VI	0.9588	32	74	0.2333	0.7667	0.9336	0.9710	0.9954	0.0409	0.0426
Model VII	0.8726	7	34	0.9309	0.0691	0.7856	0.9115	0.9763	0.1343	0.1539
Model VIII	0.2106	6	10	0.9900	0.0100	0.0646	0.1228	0.2661	0.2273	1.0794
Panel B: Box Industry N = 53										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9679	6	17	0.0942	0.9058	0.9464	0.9713	0.9957	0.0281	0.0291
Model II	0.2861	1	6	0.9999	0.0001	0.1559	0.2490	0.3020	0.2270	0.7936
Model III	0.9689	5	18	0.0942	0.9058	0.9464	0.9727	0.9971	0.0283	0.0292
Model IV	0.9803	8	24	0.0741	0.9259	0.9678	0.9867	1.0000	0.0211	0.0215
Model V	0.5788	5	8	0.9996	0.0004	0.4484	0.5524	0.6767	0.2347	0.4055
Model VI	0.9808	10	25	0.0741	0.9259	0.9678	0.9875	1.0000	0.0212	0.0216
Model VII	0.6357	7	9	0.9996	0.0004	0.4509	0.6267	0.8347	0.2562	0.4031
Model VIII	0.5859	1	7	0.7713	0.2287	0.4139	0.5239	0.7354	0.2351	0.4013
Panel C: Clothing Industry N = 160										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.8601	8	20	0.5711	0.4289	0.7969	0.8575	0.9365	0.0959	0.1115
Model II	0.6803	4	9	0.9997	0.0003	0.5217	0.6934	0.8524	0.2145	0.3153
Model III	0.8919	9	23	0.2801	0.7199	0.8345	0.8953	0.9588	0.0749	0.0840
Model IV	0.8888	15	29	0.5711	0.4289	0.8308	0.8875	0.9646	0.0909	0.1023
Model V	0.7293	7	15	0.9111	0.0889	0.5969	0.7497	0.8877	0.1971	0.2703
Model VI	0.9158	12	32	0.2397	0.7603	0.8569	0.9227	0.9725	0.0653	0.0713
Model VII	0.6996	6	17	0.9174	0.0826	0.5280	0.7129	0.8785	0.2164	0.3093
Model VIII	0.2573	3	7	0.9911	0.0089	0.0908	0.1807	0.3192	0.2445	0.9501
Panel D: Food Industry N = 218										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.8913	22	47	0.5286	0.4714	0.8424	0.9116	0.9813	0.1120	0.1256
Model II	0.9444	6	50	1.0000	0.0000	0.9298	0.9771	0.9975	0.0975	0.1033
Model III	0.9787	22	123	0.1160	0.8840	0.9688	0.9924	0.9994	0.0285	0.0291
Model IV	0.8973	22	54	0.5141	0.4859	0.8431	0.9181	0.9896	0.1094	0.1219
Model V	0.9502	8	71	1.0000	0.0000	0.9452	0.9823	0.9981	0.0949	0.0999
Model VI	0.9809	26	126	0.1155	0.8845	0.9712	0.9933	0.9998	0.0268	0.0273
Model VII	0.9449	10	92	1.0000	0.0000	0.9358	0.9747	0.9985	0.0991	0.1049
Model VIII	0.3180	4	7	0.9998	0.0002	0.1367	0.2508	0.4260	0.2426	0.7628
Panel E: Gold Industry N = 43										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.8359	11	12	0.7255	0.2745	0.7593	0.9053	0.9970	0.1966	0.2352
Model II	0.8206	4	5	0.9999	0.0001	0.7590	0.8919	0.9793	0.2298	0.2801
Model III	0.9150	15	17	0.5981	0.4019	0.9008	0.9720	1.0000	0.1353	0.1478
Model IV	0.8431	12	13	0.7255	0.2745	0.7867	0.9136	1.0000	0.1980	0.2348
Model V	0.8272	4	5	0.9999	0.0001	0.7768	0.9040	0.9793	0.2270	0.2745
Model VI	0.9158	16	17	0.5981	0.4019	0.9009	0.9720	1.0000	0.1348	0.1472

Model VII	0.8223	4	7	0.9999	0.0001	0.7731	0.8876	0.9725	0.2270	0.2761
Model VIII	0.4503	8	8	0.9367	0.0633	0.1926	0.2553	0.7375	0.3392	0.7532
Panel F: Personal Services Industry N = 47										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Model I	0.9244	13	14	0.5228	0.4772	0.9056	0.9492	1.0000	0.1051	0.1137
Model II	0.6178	6	7	0.8992	0.1008	0.3877	0.6188	0.8779	0.2788	0.4512
Model III	0.9448	16	16	0.2314	0.7686	0.9239	0.9550	1.0000	0.0609	0.0645
Model IV	0.9298	17	18	0.5228	0.4772	0.9227	0.9550	1.0000	0.1066	0.1147
Model V	0.6589	10	11	0.8826	0.1174	0.4250	0.6963	0.9744	0.2928	0.4443
Model VI	0.9502	19	20	0.2314	0.7686	0.9387	0.9664	1.0000	0.0617	0.0650
Model VII	0.6239	12	12	0.8826	0.1174	0.3817	0.5801	0.9926	0.2943	0.4718
Model VIII	0.5234	4	4	0.8238	0.1762	0.2856	0.4884	0.6731	0.2552	0.4876

Note: (a) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs that are at least 99% efficient (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (b) DEA models were the variable return of scale output-orientation. The variable return of scale, input-orientation models, generated similar results, untabulated. (c) Model I was adapted from Demerjian et al. (2012) and Demerjian et al. (2013). Model II was the transit model between Model I and Model III. Model III was adapted from Harrison and Rouse (2016). Model IV was the DuPont ratio model adapted from Feroz et al. (2001,2003). Model V was the transit model between Model IV and VI. Model VI was adapted from the stage one DEA model from Seiford and Zhu (1999) and Zhu (2000). Model VII was the funding model adapted from Smith (1990). Model VIII was adapted from the stage two DEA model from Seiford and Zhu (1999) and Zhu (2000).

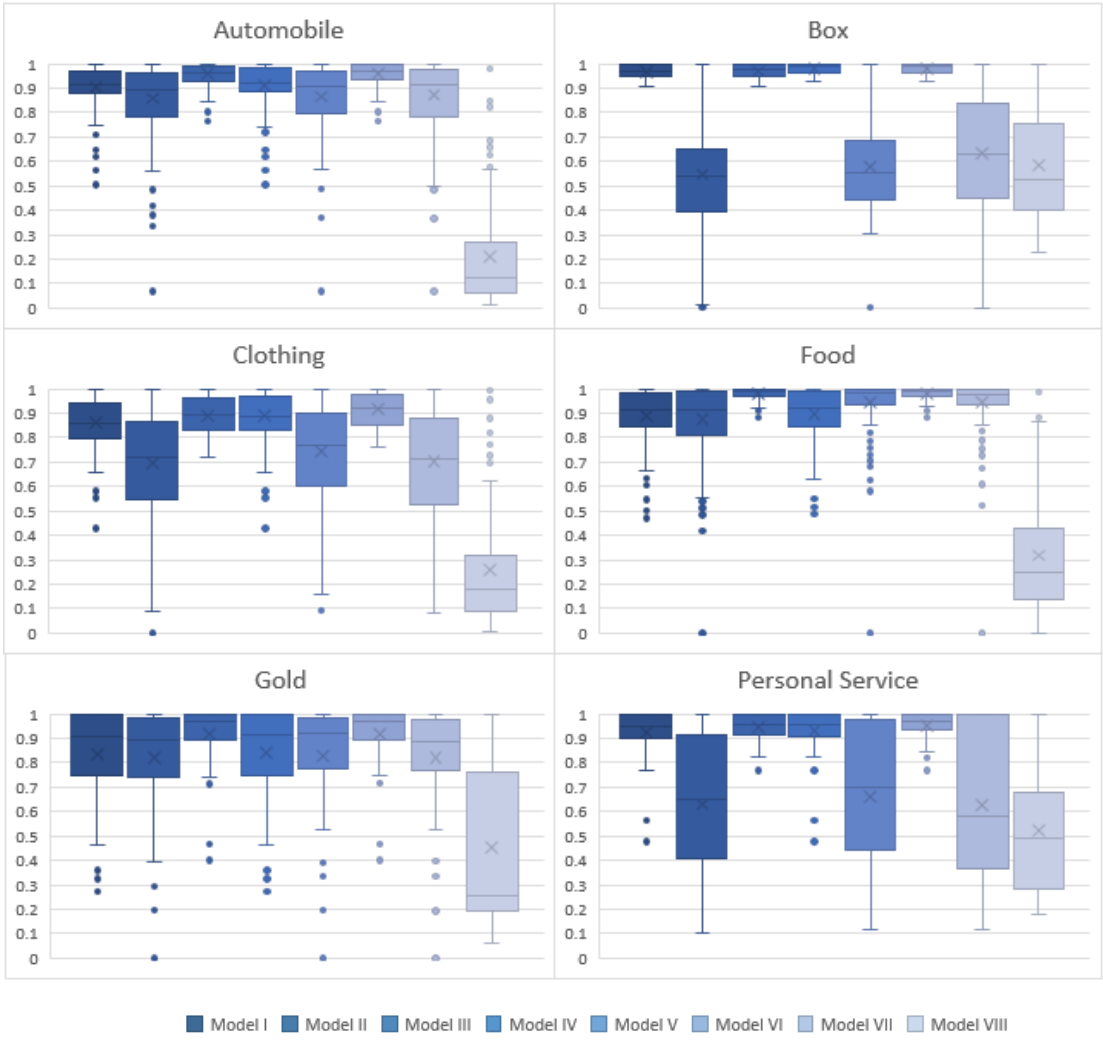


Figure 5 - 6 Test Three B Box Plot for the Distribution of Efficiency Scores

Variation in Financial DEA results – Test B

Table 5 - 22 reports Test B results, which include Spearman's ranking correlations among all models. The corresponding results in Pearson correlations can be found in Appendix 10. Test B aims to examine P_{3b} that alternative accounting variables generate diverse Financial DEA constructs and demonstrate discriminant validity where they measure different sets of indicators. Since the indicators are various, the relative domain of construct should be different in a formative model. The disturbance on the construct can result from measurement errors on individual measures that diverge from the domain of interest (Mackenzie et al., 2005).

Table 5 - 22 Test Three B Variation in Financial DEA Results (Spearman's)

Panel A: Automobile Industry N = 253

	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	**-0.1384**	1.0000						
Model III	0.5677	0.5740	1.0000					
Model IV	0.9747	-0.1960	0.5157	1.0000				
Model V	**-0.1178**	0.9359	0.5559	**-0.1123**	1.0000			
Model VI	0.6052	0.4291	0.9292	0.6318	0.5194	1.0000		
Model VII	**-0.1093**	0.8897	0.5144	**-0.1160**	0.9452	0.4608	1.0000	
Model VIII	0.3409	**0.0008**	0.4037	0.3057	**-0.0398**	0.3913	**-0.0753**	1.0000

Panel B: Box Industry N = 53

	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.4742	1.0000						
Model III	0.9436	0.5413	1.0000					
Model IV	0.8305	**0.3188**	0.8301	1.0000				
Model V	0.6074	0.8127	0.6733	0.5111	1.0000			
Model VI	0.8031	0.3551	0.8577	0.9784	0.5403	1.0000		
Model VII	0.5741	0.8025	0.6370	0.4556	0.9825	0.4825	1.0000	
Model VIII	**0.2610**	**0.2719**	**0.2119**	0.0382	0.1356	0.0191	0.1257	1.0000

Panel C: Clothing Industry N = 160

	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.1036	1.0000						
Model III	0.7379	0.5421	1.0000					
Model IV	0.8660	-0.0071	0.6125	1.0000				
Model V	0.0991	0.8736	0.4939	0.1540	1.0000			
Model VI	0.6683	0.3734	0.8701	0.7785	0.5259	1.0000		
Model VII	-0.0223	0.7721	0.3652	0.0601	0.9287	0.4294	1.0000	
Model VIII	0.4039	-0.2550	0.1665	0.3607	-0.2355	0.1813	-0.3308	1.0000

Panel D: Food Industry N = 218								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	-0.4165	1.0000						
Model III	0.3892	0.4485	1.0000					
Model IV	0.9659	-0.3972	0.3856	1.0000				
Model V	-0.1911	0.8731	0.5667	**-0.1533**	1.0000			
Model VI	0.4530	0.3516	0.9446	0.4874	0.5417	1.0000		
Model VII	**-0.0980**	0.7792	0.5982	**-0.0702**	0.9144	0.5560	1.0000	
Model VIII	0.2177	-0.2803	**-0.1442**	0.1961	**-0.1591**	**-0.1123**	-0.1413	1.0000
Panel E: Gold Industry N = 43								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	**0.2456**	1.0000						
Model III	0.8183	0.5764	1.0000					
Model IV	0.9440	**0.3154**	0.8477	1.0000				
Model V	**0.2945**	0.9890	0.6246	**0.3646**	1.0000			
Model VI	0.7946	0.5920	0.9943	0.8554	0.6396	1.0000		
Model VII	**0.2759**	0.9726	0.5966	**0.3422**	0.9856	0.6103	1.0000	
Model VIII	**0.1368**	**-0.2965**	**0.0208**	**0.0590**	**-0.2752**	**-0.0135**	**-0.2918**	1.0000
Panel F: Personal Services Industry N = 47								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.4956	1.0000						
Model III	0.8633	0.6902	1.0000					
Model IV	0.9005	0.4502	0.7591	1.0000				
Model V	0.4085	0.8924	0.5739	0.5580	1.0000			
Model VI	0.7739	0.6310	0.8932	0.8728	0.7199	1.0000		
Model VII	**0.2639**	0.7701	0.4386	0.4208	0.8833	0.5901	1.0000	
Model VIII	**-0.0603**	**0.3081**	**0.0520**	**-0.0976**	**0.2261**	**-0.0096**	**0.2549**	1.0000

Note: (a) The numbers in bold have p-values greater than 0.01. (b) The numbers in grey shadows are less than 0.2, which demonstrate discriminant validity. (c) DEA models are the variable return of scale output-orientation. The variable return of scale, input orientation models, generate similar results, untabulated. (d) Model I was adapted from Demerjian et al. (2012) and Demerjian et al. (2013). Model II was the transit model between Model I and Model III. Model III was adapted from Harrison and Rouse (2016). Model IV was the DuPont ratio model adapted from Feroz et al. (2001,2003). Model V was the transit model between Model IV and VI. Model VI was adapted from the stage one DEA model from Seiford and Zhu (1999) and Zhu (2000). Model VII was the funding model adapted from Smith (1990). Model VIII was adapted from the stage two DEA model from Seiford and Zhu (1999) and Zhu (2000). (e) The Pearson correlation results can be found in Appendix 10.

In Table 5 - 22, to demonstrate discriminant validity, the statistical correlations are expected to not correlate or correlate weakly (low correlation coefficient < 0.2, or insignificant p-values > 0.1). The numbers in bold have p-values greater than 0.01. The numbers in grey shadows are less than 0.2, which demonstrate discriminant validity. Table 5 - 22 reports Spearman's ranking correlation. The corresponding Pearson correlation results can be found in Appendix 10. The correlations reported in Table 5 - 22 are calculated using the VRS output orientation DEA results. The VRS input-orientation results (untabulated) generated very similar results.

Overall, the results confirm P_{3b} that alternative accounting variables generate diverse constructs of firm performance for different sets of indicators, showing discriminant validity. Furthermore, the magnitude of discriminant validity is affected by the level of change in indicators and industry characteristics. The key findings are threefold.

First, the correlations showing the strongest discriminant validity relate to Model VIII. Table 5 - 17 indicates that Model VIII incorporated quite different indicators in the research design than Models I to VII. In Table 5 - 22, the cells relating to Model VIII most frequently demonstrate discriminant validity across all industries. For example, in the box industry (Table 5 - 22, Panel B) and the personal services industry (Table 5 - 22, Panel F), most cells showing discriminant validity relate to Model VIII.

Second, replacing a sole indicator with another increases the discriminant validity of Financial DEA results. For example, in Table 5 - 22, Model II, V and VII, replaced *SALES*, the sole output indicator, with *NI* leading to statistically significant discriminant validity compared with the initial model. Model II is a transitional model between Model I and Model III. Model V is a transitional model between Model IV and Model VI. Recalling the model design in Table 5 - 17, Model II, V and VII only have *NI* as the output, instead of having *SALES* or *SALES* and *NI*. When *NI* is incorporated as an additional output to *SALES*, *NI* is not assigned much weighting by DEA compared to *SALES*. For instance, in the box industry, the *NI* was assigned the mean weighting of 0.0001 and 0.0003 in Model III and Model VI. By comparison, when *NI* replaced *SALES* as the sole output, the weighting is 100%. *NI* and *SALES* are different indicators, as evidenced by the accounting equation and low correlations reported in Appendix 9. As a result, replacing the output SALES with NI leads to the Financial DEA results having strong discriminant validity.

Third, an additional indicator does not impact the discriminant validity of Financial DEA results significantly. When the *NI* is the additional output to *SALES*, the Financial DEA results do not change materially. For example, Model VI compared to Model IV, and Model III compared to Model I, *NI* is the additional output to *SALES*. As evident in Table 5 - 22, the correlations between these models do not show discriminant validity. Also, the change of weightings on *NI* between models is quite marginal. For instance, the box industry has assigned a mean weight of 0.0001 to the *NI* in Model III and a mean weight of 0.0003 to *NI* in Model VI (untabulated).

Similarly, when the *CEQ* is the additional input, the Financial DEA results do not change materially. For instance, Model IV compared to Model I, and Model VI compared to Model III incorporate the *CEQ* as an additional input. As evident in Table 5 - 22, the correlations between these models do not demonstrate discriminant validity. Also, the change of weightings on *CEQ* between models was quite marginal. For instance, the automobile industry was assigned a mean weight of 0.0010 (the weightings for other variables were: *COGS* – 0.0107, *AT* – 0.0064, *XOPR* – 0.0175, *SALES* – 0.0167) to the *CEQ* in Model IV and a mean weighting of 0.0000 to the *CEQ* in Model VI (the weightings for other variables were: *XOPR* – 0.0067, *COGS* – 0.0059, *AT* – 0.0010, *SALES* – 0.0063, *NI* – 0.0001).

Fourth, the industry characteristics impact the magnitude of discriminant validity of Financial DEA results. The relatively less homogeneous industries show more evidence of discriminant validity. For example, in Table 5 - 22, the clothing, food, and gold industries have 11, 12 and 13 cells, respectively, showing discriminant validity. By contrast, the more homogeneous industry, the box industry, has the fewest cells (eight), providing evidence for discriminant validity.

To sum up, the results in Table 5 - 22 confirm P_{3b} that for different sets of indicators, alternative accounting variables define diverse constructs and demonstrate discriminant validity. Also, the replacement of sole indicators leads to stronger discriminant validity than an additional indicator. Furthermore, with a lower homogeneity level in the sample sets, the Financial DEA results show stronger discriminant validity.

Additional Analysis

Given researchers have suggested deleting DMUs with negative values rather than transforming those values, this section reports an additional test when the sample was

reduced to observations with exclusively positive values. As reported in Table 5 - 23, this section cleaned the sample further from Table 5 - 15 by deleting the observations with negative values. The negative values all relate to net income (*NI*) and common equity (*CEQ*).

As a result, the samples were refined to a smaller size. The automobile industry had 185 DMUs, the box industry had 45 DMUs, the clothing industry had 132 DMUs, the food industry had 163 DMUs, the gold industry had 23 DMUs, and the personal services industry had 29 DMUs.

Table 5 - 23 Test Three Sample without Negative Values

Industry	**Auto-mobile**	**Box**	**Clothing**	**Food**	**Gold**	**Personal Services**
Full sample for the main test	253	53	160	218	43	47
Delete observation with negative	-68	-8	-28	-55	-20	-18
Subsample for the additional test	**185**	**45**	**132**	**163**	**23**	**29**

Given that this section has more homogeneous samples, after removing the low-performing firms with negative *NI* and *CEQ*, the Financial DEA results relating to Test B are expected to be impacted. Table 5 - 24 reports the additional test results, which include Spearman's ranking correlations among all models. The corresponding results in Pearson correlations can be found in Appendix 11. The key findings are twofold.

Table 5 - 24 Test Three B with Subsamples – Variation in Financial DEA Results (Spearman's)

Panel A: Automobile Industry N = 185

	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.7593	1.0000						
Model III	0.9682	0.7956	1.0000					
Model IV	0.9569	0.6846	0.9243	1.0000				
Model V	0.7372	0.9474	0.7735	0.7332	1.0000			
Model VI	0.8974	0.7270	0.9295	0.9492	0.7894	1.0000		
Model VII	0.6980	0.9155	0.7328	0.6797	0.9567	0.7343	1.0000	
Model VIII	0.3626	0.1988	0.3426	0.3118	**0.1420**	0.2591	0.1266	1.0000

Panel B: Box Industry N = 45

	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.6390	1.0000						
Model III	1.0000	0.6390	1.0000					
Model IV	0.7420	0.3906	0.7420	1.0000				
Model V	0.5516	0.7983	0.5516	0.6040	1.0000			
Model VI	0.7420	0.3906	0.7420	1.0000	0.6040	1.0000		
Model VII	0.5926	0.8256	0.5926	0.5789	0.9866	0.5789	1.0000	
Model VIII	0.3289	0.3225	0.3289	-0.0066	0.1445	-0.0066	0.1936	1.0000

Panel C: Clothing Industry N = 132

	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.6235	1.0000						
Model III	0.9567	0.6668	1.0000					
Model IV	0.9094	0.5392	0.8774	1.0000				
Model V	0.6320	0.9275	0.6749	0.6412	1.0000			
Model VI	0.8710	0.5816	0.9127	0.9659	0.6817	1.0000		
Model VII	0.5030	0.7745	0.5432	0.5220	0.8706	0.5581	1.0000	
Model VIII	0.3880	0.3324	0.3457	0.3183	0.3099	0.2739	0.2909	1.0000

Panel D: Food Industry N = 163								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.4570	1.0000						
Model III	0.9914	0.4825	1.0000					
Model IV	0.9375	0.3939	0.9272	1.0000				
Model V	0.4433	0.9457	0.4682	0.4731	1.0000			
Model VI	0.9277	0.4286	0.9362	0.9861	0.5097	1.0000		
Model VII	0.4420	0.9089	0.4646	0.4632	0.9452	0.4933	1.0000	
Model VIII	**0.0380**	**0.1325**	**0.0413**	**0.0824**	**0.1736**	**0.0895**	**0.1078**	1.0000
Panel E: Gold Industry N = 23								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.7205	1.0000						
Model III	1.0000	0.7205	1.0000					
Model IV	1.0000	0.7205	1.0000	1.0000				
Model V	0.7786	0.9822	0.7786	0.7786	1.0000			
Model VI	1.0000	0.7205	1.0000	1.0000	0.7786	1.0000		
Model VII	0.7424	0.9498	0.7424	0.7424	0.9686	0.7424	1.0000	
Model VIII	**0.0000**	**-0.2445**	**0.0000**	**0.0000**	**-0.1827**	**0.0000**	**-0.1921**	1.0000
Panel F: Personal Services Industry N = 29								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.4955	1.0000						
Model III	0.9516	0.5922	1.0000					
Model IV	0.9896	0.5235	0.9616	1.0000				
Model V	0.6105	0.9079	0.6957	0.6326	1.0000			
Model VI	0.9516	0.5922	1.0000	0.9616	0.6957	1.0000		
Model VII	0.7315	0.6766	0.7672	0.7480	0.7676	0.7672	1.0000	
Model VIII	**0.1035**	**-0.0346**	**0.1259**	**0.1216**	**-0.1806**	**0.1259**	**-0.0691**	1.0000

Note: (a) The numbers in bold have p-values greater than 0.01. (b) The numbers in grey shadows are less than 0.2, which demonstrate discriminant validity. (c) DEA models are the variable return of scale output-orientation. The variable return of scale, input-orientation models, generate similar results, untabulated. (d) Model I was adapted from Demerjian et al. (2012) and Demerjian et al. (2013). Model II was the transit model between Model I and Model III. Model III was adapted from Harrison and Rouse (2016). Model IV was the DuPont ratio model adapted from Feroz et al. (2001,2003). Model V was the transit model between Model IV and VI. Model VI was adapted from the stage one DEA model from Seiford and Zhu (1999) and Zhu (2000). Model VII was the funding model adapted from Smith (1990). Model VIII was adapted from the stage two DEA model from Seiford and Zhu (1999) and Zhu (2000). (e) The Pearson correlation results can be found in Appendix 11.

First, as reported in Table 5 - 24, the cells demonstrating discriminant validity relate to Model VIII only. Unlike Table 5 - 22, the result of Test B with the full sample, in Table 5 - 24, the results of Model VII do not diverge from the other models (do not demonstrate discriminant validity) after removing the poor performing firms (i.e. DMUs with negative *NI* and *CEQ*). Considering the indicators in each model, Model VII is comparable to Model I to VI in that *CEQ* is one of the inputs and *NI* is one of the outputs. Further, the two variables were refined to be more homogeneous after deleting the negative values. According to the accounting equation, although Model VII incorporated liability as one of the inputs, the liability is the difference between assets and equity. The three indicators are all significantly and positively correlated, as shown in Appendix 9. Therefore, it is not surprising that Model VII lost its discriminant validity compared with Model I to VI after the refinement of the sample. By comparison, Model VIII used entirely different indicators from Model I to VII. The Financial DEA model used the *NI* and *SALES* as inputs to generate *MKVALT*.

Second, the industry features still impact the Financial DEA results, but to a lesser magnitude. The relatively heterogeneous industries, such as the food, gold, and personal services industries, retain the strong discriminant validity of Model VIII against the other models from Table 5 - 22. The relatively homogeneous industries, such as the box industry, lose the discriminant validity in Models I to VII and partially in Model VIII due to the deletion of poor performing firms.

In summary, the additional test further refined the sample with the deletion of the observations with negative *NI* and *CEQ*. The further refinement leads to more homogeneous samples since the relatively poor performing firms were removed. Test B results show a much clearer trend in that only Model VIII shows discriminant validity against the other models. The homogeneity level within industries still impacts the magnitude of the discriminant validity, but to a much lower level than in the full samples.

5.4.2. Summary

This test examines the impact on the construct validity of using Financial DEA models with different specifications. It found that the research context influences convergent validity and discriminant validity across different Financial DEA models. When models are constructed using similar indicators, Financial DEA results demonstrated convergent validity among models. The more homogeneous the samples are, the stronger the convergent validity

identified in the Financial DEA results. When models are constructed using different indicators, Financial DEA results had discriminant validity among models. The more homogeneous the samples are, the weaker the discriminant validity identified in the Financial DEA results.

5.5. Chapter Summary

This chapter examined the empirical impact of selective measurement errors on Financial DEA. Three tests illustrated selective sources of measurement errors in various research scenarios.

The first test examined the impact of price factors on the results of Financial DEA. The findings show the more variation in the price factor leads to more divergent Financial DEA results from conventional DEA, which measures technical efficiency using physical measures.

The second test examined the impact of alternative stock and flow forms of accounting variables on the results of Financial DEA. This test illustrated a research context where capital was measured by three alternative accounting variables: the *GPPE*, *NPPE*, and *DP*. The findings demonstrated that alternative stock and flow forms of accounting variables contain various factors, such as accounting choices and operational information. The variation of these factors influences Financial DEA results to various degrees. The more variation of accounting choices and operational information the more the Financial DEA results diverged from the comparison point, where no accounting choices or operational information were captured by Financial DEA.

The third test examined the impact of alternative accounting variables on the results of Financial DEA. The test illustrated six industry contexts and two sets of Financial DEA models, covering 12 models in total. The findings demonstrated that industry features could impact the construct validity of the Financial DEA. When Financial DEA models had the same set of indicators, results demonstrated convergent validity. The more homogeneous the sample was, the stronger the convergent validity. When Financial DEA models had different sets of indicators, results demonstrated discriminant validity. The more homogeneous the sample was, the weaker the discriminant validity.

Chapter 6: Discussion

6.1. Chapter Introduction

Researchers have been using Financial DEA models with different views, and thus there is not a clearly articulated conceptual basis. A conceptual foundation is needed to clarify the nature of different Financial DEA models and the type of performance they are measuring. By investigating the hybrid nature of Financial DEA, this study proposes a conceptual foundation for Financial DEA, which can serve as a map to guide Financial DEA applications.

To answer the overarching goal of the study, which was to develop a conceptual foundation for Financial DEA, this chapter synthesises the findings of the three research questions investigated in Chapters 4 and 5. The three research questions are:

> ***RQ1.*** *What dimensions of firm performance do Financial DEA models measure, and how have they been used?*
>
> ***RQ2.*** *What are the methodological issues when applying Financial DEA?*
>
> ***RQ3.*** *What are the empirical impacts of methodological choices on the results of Financial DEA?*

Chapter 6 is structured as follows. Section 6.2 discusses the results of *RQ1* by synthesising the dimensional firm performance constructs found in Chapter 4 (section 4.2.2) with Porter's value chain framework discussed in Chapter 2 (section 2.4.2). This section discusses Financial DEA studies in terms of firm performance. This section also discusses the interrelationships between the dimensional constructs used by researchers in the context of business structure. The interrelationships between constructs identify the sources of methodological issues in Financial DEA research. Section 6.3 discusses the results of *RQ2* by synthesising the methodological issues found in Chapter 4 (section 4.3) with the alternative lenses of measurement models (section 2.5). This section continues the discussion of *RQ1* to interpret various sources of methodological issues of Financial DEA in the context of business structure. Section 6.4 synthesises the methodological issues illustrated empirically in *RQ3* (Chapter 5) with the previous two sections (section 6.2 and section 6.3) to provide

quantitative evidence for the conceptual foundation of Financial DEA. Section 6.5 summarises Chapter 6 and synthesises the findings of the three research questions above to provide a comprehensive conceptual foundation facilitating the application of Financial DEA.

6.2. Research Question One

This section discusses the finding of *RQ1:*

> ***RQ1.*** *What dimensions of firm performance do Financial DEA models measure, and how have they been used?*

This section synthesises the findings of the literature typology in Chapter 4 (section 4.2) with Porter's value chain framework reviewed in Chapter 2 (section 2.4.2) to provide a conceptual framework at the construct level for Financial DEA. This section synthesises the dimensional constructs of firm performance measured by Financial DEA in the context of business structure, where the interaction among various levels of the business structure is considered. This section organises the dimensional constructs of firm performance into a nested framework, starting with the physical production level, aggregating at each level, and finishing at the market level.

In section 6.2.1, the dimensional constructs identified by the typology in Chapter 4 (section4.2) are organised to reflect the business structure and provide a conceptual framework at the dimensional construct level. In section 6.2.2, the dimensional constructs from the typology of Chapter 4 (section 4.2) are synthesised with Porter's value chain framework discussed in Chapter 2 (section 2.4.2) to summarise the current state of Financial DEA studies. Section 6.2.3 summarises the section.

6.2.1. A Conceptual Framework for Financial DEA at the Construct Level

In Chapter 4 (section 4.2.1), 12 dimensional firm constructs were identified from the Financial DEA literature typology. These dimensional constructs relate to various activities within the firm. Among the various views of firm performance, the value chain framework by Porter (1985) seems to best explain these dimensional constructs in a business structure. As illustrated in Figure 6 - 1, the dimensional constructs are identified at various levels in a business structure. The structure begins at the physical production level, which aligns with the origin of DEA in economics. It finishes at the share market level, which is the highest

level of aggregation in terms of the business structure identified in the Financial DEA literature. The business structure aggregates and expands to a higher level for every step, distancing away from the physical production process.

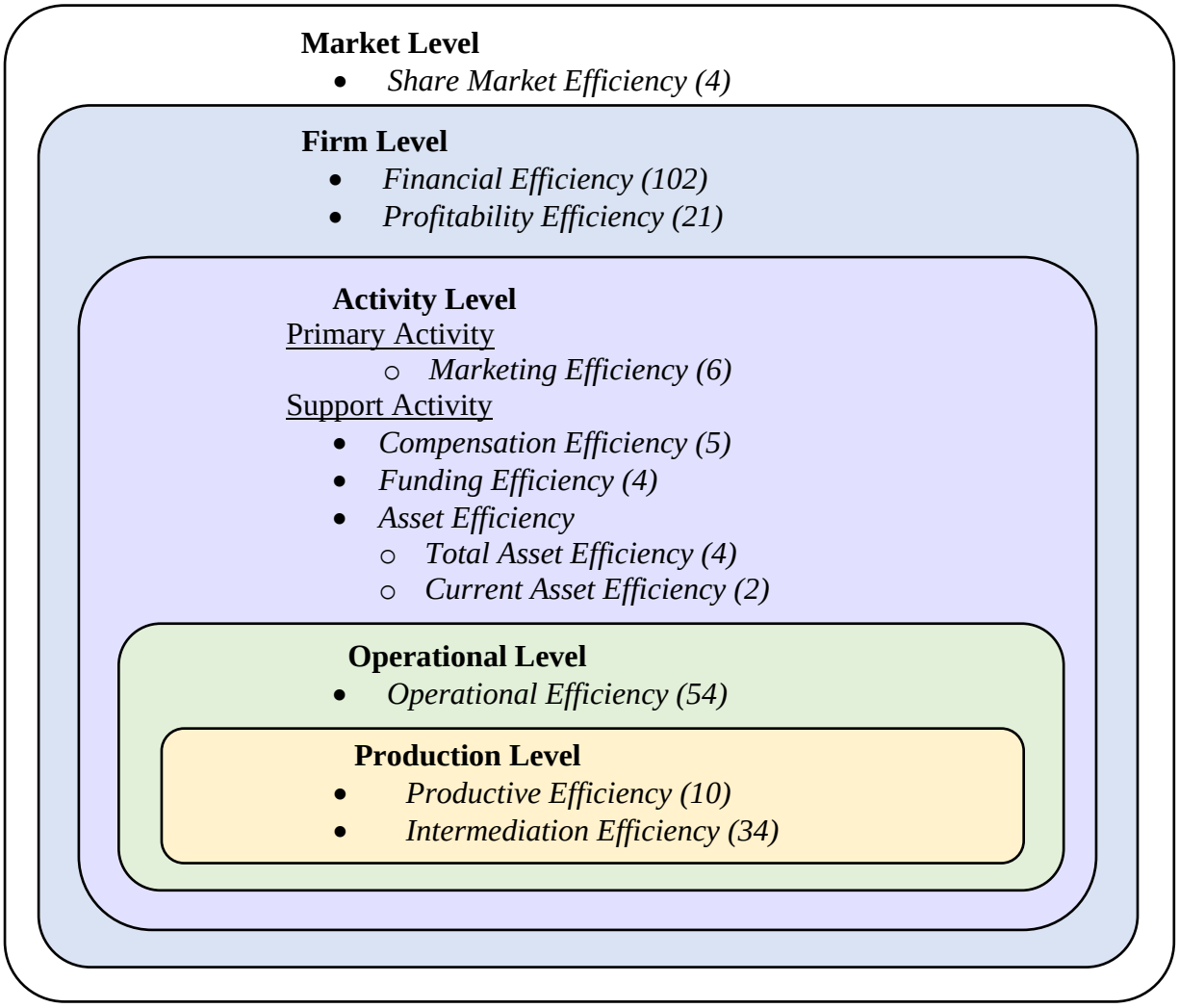


Figure 6 - 1 The Conceptual Framework for Financial DEA at the Construct Level
Note. the numbers in the brackets are the number of models identified in the literature.

At the production level (yellow box, Figure 6 - 1), Financial DEA measures the productive efficiency of a physical production process. This productive efficiency follows the definition of technical efficiency by Farrell (1957). In conventional DEA, the inputs and outputs are physical measures. For example, in banks, the amount of labour and capital generates the number of transactions and documents processed (Paradi et al., 2011). In Financial DEA, the accounting measures closely proxy physical production elements. For example, the

production efficiency in the banking sector uses the cost of labour and capital to produce the value of the transactions (Kaffash et al., 2018; Oral & Yolalan, 1990). This type of Financial DEA model is relatively rare in the literature, given the difficulties in collecting financial data at this level of operation. In Chapter 4 (section 4.2.1), there were ten studies identified in this category.

Intermediation efficiency is another type of efficiency at the production level (yellow box, Figure 6 - 1) due to the monetary nature of most products in banks. As intermediates, banks make loans and investments based on the monetary assets they gather (funds) (Paradi et al., 2011). In intermediation efficiency models, accounting variables are the direct measure of production elements. This type of Financial DEA model was only noted in the literature related to the banking sector. In Chapter 4 (section 4.2.1), there were 34 studies identified in this category.

Financial DEA measures the dimensional construct of operational efficiency at the operational level (green box, Figure 6 - 1), which essentially aggregates multiple production processes at the production level (yellow box, Figure 6 - 1). Operational efficiency is named after the "operational activities" in Porter's value chain (section 2.4.2) that converts the inputs to outputs (Porter, 1985). However, measurement of the productive efficiency of a specific production process using Financial DEA is rare due to the availability of accounting data and various processes. Also, DEA is a black box in nature in that it does not need to know the process by which DMUs become more efficient (Rouse et al., 2010). By comparison, researchers mostly measured aggregated production processes at the level of a firm within industries. In the typology in Chapter 4 (section 4.2.1), 54 studies fit into this category. For example, researchers measured the technical efficiency of dairy-manufacturing firms (Aparicio & Kapelko, 2018; Kapelko & Lansink, 2017). The Financial DEA models mainly used labour cost, the cost of material and fixed assets as inputs, and the revenues as the output. However, dairy-manufacturing firms can cover various activities such as processing raw milk and producing a range of milk products such as liquid milk, frozen milk products, and fermented milk products (Minj et al., 2020). The Financial DEA models measured the operational efficiency of dairy-manufacturing firms without further distinguishing the differences between production lines. The expansion from the production level (yellow box, Figure 6 - 1) to the operational level (green box, Figure 6 - 1) reaches a higher aggregation of the Financial DEA model. However, the higher aggregated model loses

detailed information on specific production processes. The heterogeneity between various production lines is aggregated into accounting variables.

At the activity level (purple box, Figure 6 - 1), the focus expands from the operational activities to other activities, including “primary activities” and “support activities” defined in Porter’s value chain. The primary activities are defined as the activities that add value to the product, and the operational activity is one of the primary activities (Porter, 1985). The support activities are defined as the activities that make the primary activities more efficient (Porter, 1985).

Among all the primary activities proposed by Porter (1985), the Financial DEA studies reviewed only covered marketing efficiency, which was defined as the efficiency to minimise the advertising inputs while maximising the outputs, which convert ultimately into revenues (Rahman et al., 2019). The Financial DEA models generally use advertising expenses as the input and sales revenues as the output (Akdeniz et al., 2010; Rahman et al., 2018). However, this category is not common in the Financial DEA studies. Only six studies were identified from Chapter 4 (section 4.2.1).

Similarly, only selective support activities were covered by the Financial DEA literature. First, compensation efficiency was defined as the efficiency of compensating employees. The Financial DEA literature reviewed only had five studies for this type of efficiency, and all the studies focused on the executives’ compensation efficiency. Second, funding efficiency was defined as the incremental efficiency that can benefit the firm from the mix of capital sources. The Financial DEA models reviewed generally used the liabilities and equity as inputs and the revenue or income as the output (Oberholzer, 2014b; Oberholzer et al., 2017; Smith, 1990). In Chapter 4 (section 4.2.1), there were four studies identified in this category. Third, asset efficiency was defined as the incremental efficiency that can benefit the firm from owning assets. In the Financial DEA literature, researchers further classify the asset efficiencies as total asset efficiency (Kwon et al., 2008; Min, H. & Joo, 2009) and current asset efficiency (Joo et al., 2011) according to the timeline. There were six studies identified in this category in Chapter 4 (section 4.2.1), four studies measured the efficiency of the total asset, and two studies measured the current asset efficiency.

At the firm level, according to Porter (1985), the support activities help make primary activities more efficient and benefit the firm as a whole (blue box, Figure 6 - 1). The

aggregation of production level (yellow box, Figure 6 - 1), operational level (green box, Figure 6 - 1), and activity level (purple box, Figure 6 - 1) into the firm level (blue box, Figure 6 - 1) considers the synergy between all the activities within a firm. Compared with the separate activities, the combined firm performance at the firm level views a firm as an organically integrated organisation. However, on the downside, the integrated internal firm performance may not identify the strengths and weaknesses within a specific activity.

In the Financial DEA literature, the internal firm performance was usually represented by financial efficiency. It was defined as the efficiency of minimising financial resources while maximising the financial outputs (Demerjian et al., 2012; Demerjian et al., 2013). The seminal work by Demerjian et al. (2012) provided an example of the Financial DEA model at the firm level that incorporated a range of financial resources as inputs and the sales revenue as the output. This dimensional construct is the most common type identified from Chapter 4 (section 4.2.1). There were 102 studies in this category. Profit efficiency further narrows down the financial efficiency to generating profits from expenses (Kao & Hwang, 2008). In the typology in Chapter 4 (section 4.2.1), 21 studies fit this category,

The final category aggregates the view from the firm level (blue box, Figure 6 - 1) to the market level (white box, Figure 6 - 1). In Financial DEA, share market efficiency is the most common parameter used to measure external firm performance. Share market efficiency was defined as a firm's performance in the share market (Seiford & Zhu, 1999; Zhu, 2000). In Chapter 4 (section 4.2.1), four studies were identified in this category.[40]

In sum, Figure 6 - 1 synthesises the 12 dimensional constructs identified in Chapter 4 (section 4.2.1) with business structure and provides a nested conceptual framework of the Financial DEA at the construct level. This conceptual framework can be used to structure the literature. Researchers can use this conceptual framework to locate the level of detail to measure firm performance.

6.2.2. Evaluation of Financial DEA Literature Using Porter's Value Chain

The dimensional constructs identified from Chapter 4 (section 4.2.1) reflect Porter's value chain framework (section 2.4.2). The dimensions covered by Porter's value chain are

[40] In the Financial DEA literature typology, two models were designed to meet the performance goals of local government. This framework at the construct level did not incorporate this category since it does not belong to the for-profit firms and only a small quantity of studies was identified.

operations activities, marketing and sales activities, firm infrastructure, human resources, and technology.

The primary activities, operations activities, and marketing and sales activities were identified in the Financial DEA literature review. The measurement of operational efficiency is relatively frequent (54 models). By comparison, only a small number of studies measured funding efficiency (4 models). Three additional dimensions were not specifically covered by the Financial DEA researchers: inbound logistics, outbound logistics, and service.

Financial DEA researchers have intensively measured the efficiency of operations, reflecting DEA's origins in the measurement of productive efficiencies. Incorporating inbound and outbound activities could expand the Financial DEA measure of the production process. It could include pre- and post-manufacture/retail and extend the operating efficiency of Financial DEA studies to a more comprehensive view of the efficiency of primary activities. Similarly, the customer service activity was not specifically covered by the Financial DEA literature. The main reason is likely to be the lack of available data for the service processes. The after-sale processes may be difficult to quantify into accounting measures since not all after-sale services involve charging customers. Also, the inputs can be heterogeneous due to the variety of after-sale activities. However, by incorporating after-sale service, the measurement of firm performance would incorporate more of the primary activities driving value in businesses.

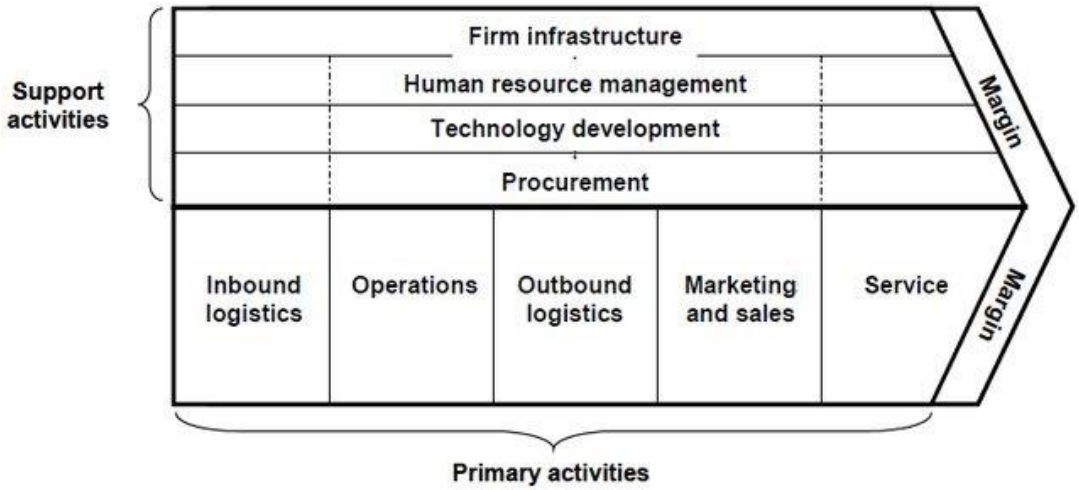


Figure 6 - 2 Porter's Value Chain
from Porter (1985)

For support activities, firm infrastructure, human resource management, and technology were covered by the Financial DEA literature reviewed. For example, accounting measurements of technology development were often used in the internal financial efficiency models (Demerjian et al., 2012). The main obstacle of measuring technical development efficiency is likely to be data availability. In the empirical test in section 5.4, most sample firms did not have separate disclosures of research and development expenses. Although the data on the intangible assets were usually available, they did not necessarily measure all intangible property used in a firm. Measuring the technical development efficiency alone, researchers could quantify the potential of future technology development. Similarly, for firm infrastructure, only the funding activities were measured in the literature reviewed; for human resource management, only the executives' compensation was studied. The Financial DEA studies did not cover procurement efficiency as other missing primary activities, which is likely to reflect data availability issues.

6.2.3. Summary

In sum, this section answers *RQ1* by synthesising the dimensional constructs found in Chapter 4 (section 4.2.1) with Porter's value chain frameworks reviewed in Chapter 2 (section 2.4.2). This section provides a conceptual framework for Financial DEA at the construct level based on Porter's value chain to discuss the interrelationship between the dimensional constructs. This section structures the literature of Financial DEA studies within the context of a business structure and reviews the Financial DEA literature in business structure.

6.3. Research Question Two

This section discusses the findings of *RQ2*:

RQ2. *What are the methodological issues when applying Financial DEA?*

This section synthesises the finding from Chapter 4 (section 4.3) with the measurement models reviewed in Chapter 2 (section 2.5) to further develop the conceptual framework for the Financial DEA modelling process. In section 6.3.1, based on the measurement models in Chapter 2, a continuum between the reflective and formative constructs is added to the framework of dimensional constructs discussed in *RQ1*. In section 6.3.2, the measurement

models in Chapter 2 are integrated into a four-quadrant framework to classify the modelling processes in Financial DEA research. In section 6.3.3, measurement errors within each quadrant are discussed in the context of Financial DEA research. Section 6.3.4 summarises this section.

6.3.1. Continuum of Measurement Models in Financial DEA

6.3.1.1. Relationships between variables and constructs

A continuum of measurement models is built based on the relationship between the dimensional constructs identified from *RQ1* and the measurement models discussed in Chapter 2 (section 2.5). There are two types of measurement models explaining the relationship between accounting variables and dimensional constructs. These are reflective models and formative models.

Reflective models tend to incorporate the accounting variables that closely reflect the underlying physical production processes. In this view, the Financial DEA and the conventional DEA are quite similar in that they are closely related to economic production theories. Therefore, one end of the continuum is the reflective model (orange end of the arrow in Figure 6 - 3), where accounting variables are selected to proxy production elements in the underlying physical production process. The reflective end covers mostly the dimensional constructs at the production level (yellow box, Figure 6 - 3) and the operational level (green box, Figure 6 - 3). For example, Aparicio and Kapelko (2018) measured the operational efficiency of dairy-manufacturing firms by using material cost, labour cost, and fixed assets as inputs to generate revenues as the output. The input accounting variables selected directly proxy the key production elements as suggested by production theories. The output, revenue, was used to proxy the output of the operation process. It has been argued that revenue can be improved to be a more accurate measure of the output volume by adjusting inventory levels (Harrison & Rouse, 2016). However, revenue is a relatively good measure considering the data availability.

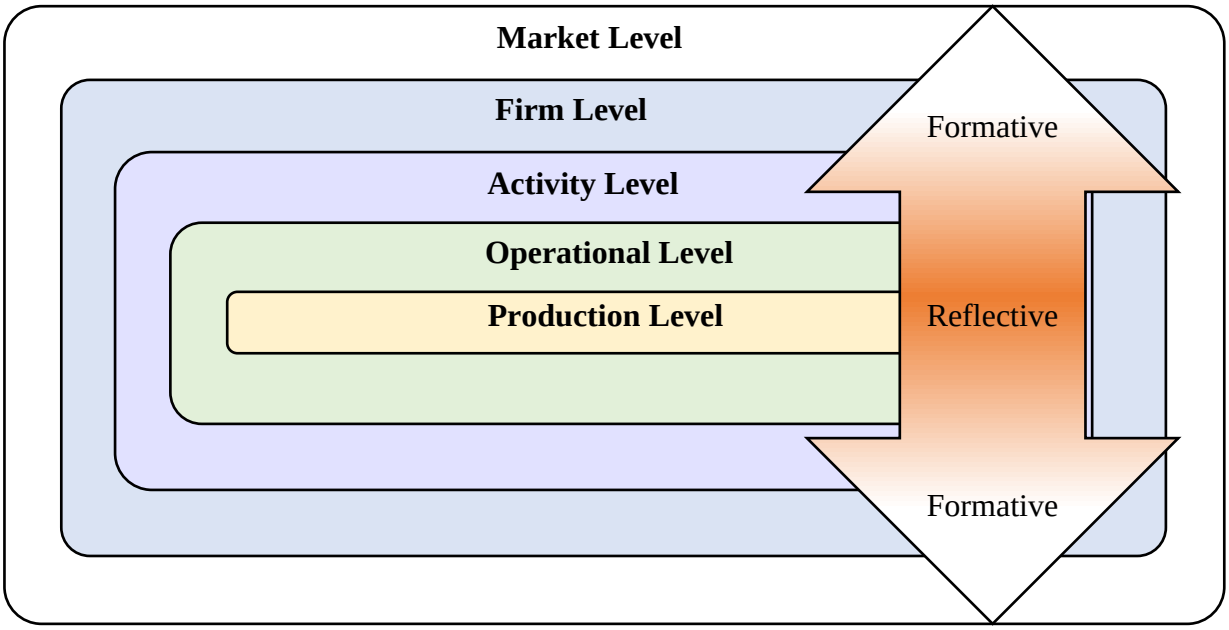


Figure 6 - 3 A Continuum of the Dimensional Constructs in Financial DEA

Formative models tend to incorporate accounting variables that are less closely related to the physical production processes. Therefore, the other end of the continuum is the formative model (white ends of the arrow in Figure 6 - 3), where accounting variables are selected at the firm level for benchmarking purposes. The link between accounting variables and specific production elements may not be direct. Rather, the inputs are the economic resources that the firm aims to use minimally. The outputs are the economic benefits that the firm aim to maximise. For example, Demerjian et al. (2012) and Demerjian et al. (2013) measured the financial efficiency at the firm level. On the input side, the Financial DEA model used seven accounting variables representing economic resources: net property plant and equipment, net operating lease, net capitalised research and development expenses, the purchased goodwill, other intangibles, the cost of goods sold, and selling, general and administrative expenses. On the output side, revenue was the output used to measure the economic gain.

6.3.1.2. Relationships between constructs and phenomenon

In the relationships between constructs and phenomenon, as discussed in section 2.5.2.3, the distinction between the reflective and formative models lies in the difference between the generality or specificity of researchers' theoretical interests (Mackenzie et al., 2005). Researchers only need general measurement models without distinguishing dimensions when

the firm performance phenomenon is not the primary interest. However, researchers need to draw specific conceptual distinctions between each dimension when the firm performance phenomenon is the primary focus of a study (Mackenzie et al., 2005; Miller et al., 2013).

In the Financial DEA context, as discussed in Chapter 2 (section 2.4), three approaches are used to specify the firm performance phenomenon: (a) the latent multidimensional approach, (b) the separate approach, and (c) the aggregated approach, and.

First, the latent multidimensional approach fits the reflective model, which assumes that the firm performance phenomenon exists deeper than its dimensions. This approach is applied when firm performance is not the primary focus of a study. For example, Demerjian et al. (2013) aimed to measure the relationship between managerial ability and quality of earnings. The firm performance phenomenon was only used to measure managerial ability, and the authors stated that they measured the firm performance phenomenon without specifying the meaning explicitly. Instead, the Financial DEA models incorporated seven inputs and one output measuring the dimensional construct of financial efficiency. In practice, researchers focus on an abstract general conceptualisation of firm performance with an implicit meaning. Researchers do not need to describe different performance dimensions (Miller et al., 2013). In Financial DEA, when the relationship between the phenomenon and dimensional construct is reflective, the relationship between dimensional constructs and accounting variables could be either reflective or formative.

Second, the separate approach fits the formative model, which treats the firm performance phenomenon as a domain of separate constructs by specifying specific constructs. This approach is used when the firm performance phenomenon is the primary focus of a study. For example, Frijns et al. (2012) examined the relationship between firm efficiency and stock returns. The authors specified firm performance as distinguished dimensions, financial efficiency, and share market efficiency. Under this approach, firm performance represents distinct constructs (Miller et al., 2013). In Financial DEA, when the relationship between the phenomenon and the dimensional construct is formative, the relationship between the dimensional construct and accounting variables could be either reflective or formative.

Third, the aggregated approach explicitly treats firm performance as a phenomenon formed by the mathematical (DEA model in the context of this study) combination of dimensions (Miller et al., 2013). As Rouse et al. (2002) and Rouse (2006) illustrated, the DEA model

could be used to integrate different performance dimensions into one composite score based on a balanced scorecard. Similarly, in the field of corporate social responsibility, DEA has been used to aggregate indicators from the three sustainability performance dimensions (economic, environmental, and social) to develop a composite index (Zharfpeykan, 2017). However, no Financial DEA research literature reviewed fits this category. The closest examples are by Seiford and Zhu (1999) and Zhu (2000). In these examples, the firm performance measured was a combination of share market efficiency and profitability efficiency (Seiford & Zhu, 1999; Zhu, 2000). The authors used a two-stage DEA model to combine these two dimensions into a composite score. However, the number of employees was measured by physical numbers rather than accounting measures, and the first-stage DEA model was not classified as Financial DEA by this study. In the future, researchers could explore incorporating varied accounting data representing different dimensions into Financial DEA models.

6.3.2. Four-quadrants of Measurement Models in Financial DEA

This section continues from section 6.3.1 to discuss the continuum of Financial DEA models and combines them with the measurement models discussed in section 2.5. These categories of measurement models can be synthesised into a four-quadrant framework for the process of Financial DEA modelling. Figure 6 - 4 illustrates four quadrants of measurement models in Financial DEA research. For each quadrant, a Financial DEA model is shown as an illustrative example in Table 6 - 1.

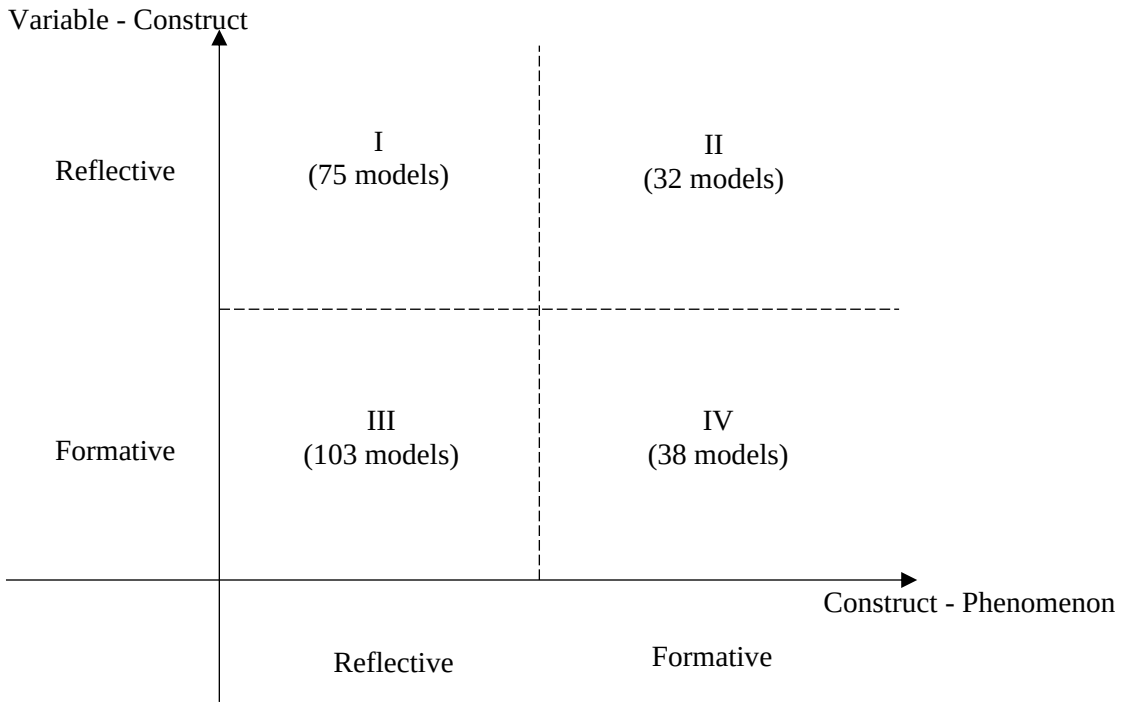
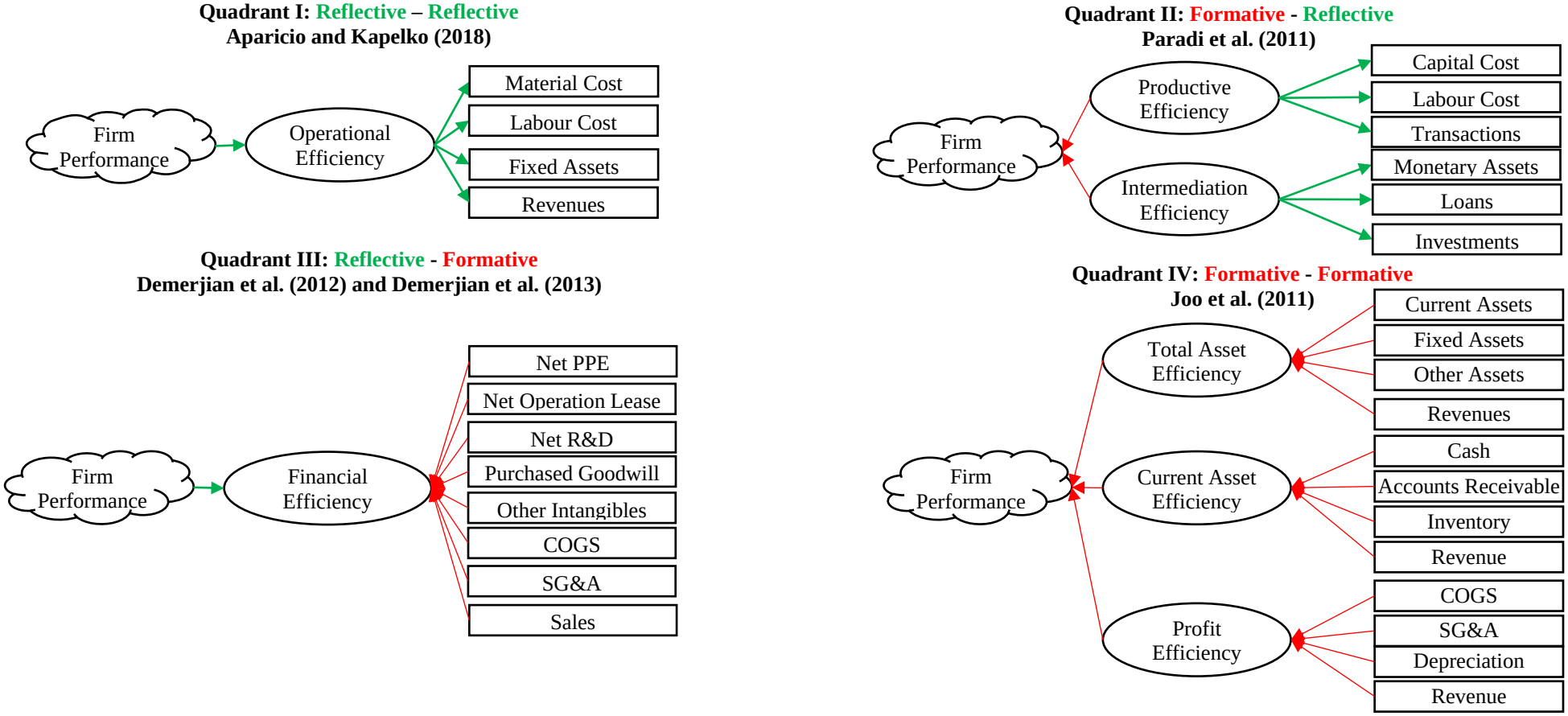


Figure 6 - 4 Four Quadrants in Financial DEA

Quadrant I incorporates two reflective models in the relationship between phenomena and constructs, and constructs and variables. In the relationship between constructs and variables, variables reflect Financial DEA dimensional constructs, which further reflect the firm performance phenomenon. For conceptualisation, firm performance is not the primary research interest, and it is treated as a latent multidimensional construct. For operationalisation, the Financial DEA model closely reflects physical productive efficiency. For example, as illustrated in Table 6 - 1, Aparico and Kapelko (2018) used the cost of material and labour, and fixed assets to produce revenues. These accounting variables closely reflect the underlying production processes of dairy manufacturers. The operational efficiency construct further reflects firm performance at a more abstract level. Since firm performance is not the primary focus in this study, researchers only need to focus on an abstract and general conceptualisation of firm performance. Researchers are not required to specify different dimensions.

Table 6 - 1 Financial DEA Examples for Measurement Models



Note: (a) The name of each quadrant follows the order of first, the relationship between the phenomena and the dimensional construct; second, the relationship between dimensional constructs and accounting variables. For example, Quadrant II is named formative-reflective, which represents, first, the relationship between the phenomena and the dimensional constructs is formative; second, the relationship between the dimensional construct and the accounting variables is reflective. (b) The red-coloured arrow represents the formative relationship, and the green-coloured arrow represents the reflective relationship. (c) The cloud shape represents a phenomenon, the circle shape represents a dimensional construct, and the rectangle shape represents an accounting variable.

Quadrant II incorporates a reflective model to explain the relationship between variables and constructs and a formative model to explain the relationship between constructs and phenomena. Between variables and constructs, different Financial DEA models reflect distinct dimensional constructs based on various production processes. These dimensional constructs further form the composite construct, the firm performance phenomenon, in the relationship between constructs and phenomena. For conceptualisation, firm performance is the primary research interest, and it is treated as a set of separate constructs. For operationalisation, the Financial DEA model closely reflects physical production processes. For example, as illustrated in Table 6 - 1, Paradi et al. (2011) measured two dimensions with different production processes using reflective models: productive efficiency and intermediation efficiency with different sets of variables. The productive efficiency measured the process of transferring the cost of capital and labour to produce transactions. The intermediation efficiency measured the process of transferring the monetary assets to loans and investments. These dimensional constructs represented different dimensions of the firm performance phenomenon since the underlying production processes were distinguished. Therefore, these dimensional constructs formed a multidimensional construct – the firm performance phenomenon.

Quadrant III incorporates a formative model in the relationship between variables and constructs and a reflective model in the relationship between constructs and phenomena. In the relationship between variables and constructs, accounting variables form composite dimensional constructs, which reflect the underlying latent construct – the firm performance phenomenon. For conceptualisation, firm performance is not the primary interest variable, and the construct is treated as a latent multidimensional construct. For operationalisation, the Financial DEA model is mainly for benchmarking purposes, which measures production processes at a relatively distant and aggregated level. For example, as illustrated in Table 6 - 1, Demerjian et al. (2012) and Demerjian et al. (2013) measured financial efficiency at the firm level. The Financial DEA model used seven accounting variables representing the economic resources at a firm level to produce revenues. The seven accounting inputs were: net property plant and equipment, net operation lease, net capitalised research and development expenses, the purchased goodwill, other intangibles, the cost of goods sold, and selling general and administrative expenses. The dimensional construct benchmarked financial efficiency at the firm level. The link between these variables and a specific production process is not direct. This dimensional construct later reflected the firm

performance phenomenon at a more abstract level. In this case, firm performance was not of primary interest. The researchers treated the firm performance as a latent multidimensional construct that did not need to be further distinguished from its various dimensions.

Quadrant IV incorporates two formative models in the relationship between variables and constructs and constructs and phenomena. The accounting variables form the domain of dimensional constructs, which form the domain of the phenomenon – firm performance. For conceptualisation, firm performance is of primary research interest, and it is the combination of separate dimensional constructs. For operationalisation, the Financial DEA model is mainly for benchmarking purposes, which measures production processes at a relatively aggregated level. For example, as illustrated in Table 6 - 1, Joo et al. (2011) formed the firm performance phenomenon with three dimensional constructs, and different combinations of accounting variables formed each dimension. The total asset efficiency used current assets and fixed assets as the input and revenues as the output. The current asset efficiency used cash, account receivables, and inventories as inputs and revenues as the output. The profitability efficiency used the cost of goods sold, selling general and administrative expenses, and the depreciation expenses as the inputs and revenue as the output. For the conceptualisation, the authors specified the meaning of the firm performance phenomenon as the combined domain of these three dimensional constructs.

In a measurement model, some of the indicators may be reflective, while others are formative. This type of model is named the mixed indicator model (Bollen, 1989; Bollen & Ting, 2000). However, this type of model is not discussed in this study since little empirical Financial DEA literature has been identified so far that belong in this category. For example, there were three common Financial DEA models in the banking sector: productive efficiency, intermediation efficiency, and profitability efficiency (Paradi et al., 2011). As stated in Chapter 4 (section 4.2.1) and section 6.2, both productive efficiency and intermediation efficiency are closer to the reflective end since the accounting variables closely relate to physical production processes. The profitability efficiency, by comparison, is closer to the formative end since the accounting variables only relate to production processes at an aggregated firm level.

6.3.3. Measurement Errors in Financial DEA

This section discusses the measurement errors in Financial DEA. This section develops the four-quadrant measurement models in Financial DEA (section 6.3.2) and synthesises them with Financial DEA methodological issues (section 4.3).

In a reflective model of the relationship between constructs and phenomena in the conceptualisation phase (Quadrant I and III in Figure 6 - 4), the phenomenon of firm performance is reflected in a dimensional construct measured by Financial DEA. The domain of the firm performance phenomenon that is not covered by the domain of Financial DEA introduces random measurement errors into the dimensional construct. For example, if firm performance intended to measure is defined as the domain of operational efficiency in Figure 6 - 1, but the dimensional construct actually measured by Financial DEA covers the domain of financial efficiency (Figure 6 - 1), the random measurement errors are mainly due to the other activities within the financial efficiency but not in the operational performance. These activities can include the efficiencies of marketing, inbound, outbound, and after-sale services. However, the interpretation of Financial DEA results is still related to the domain of operational efficiency. The remedy is to measure the firm performance phenomenon with the dimensional constructs that share the same domain, for instance, by incorporating variables representing the performance of all activities to expand the domain of the dimensional construct from operational efficiency to financial efficiency.

In a formative model of the relationship between constructs and phenomena in the conceptualisation phase (Quadrant II and IV in Figure 6 - 4), firm performance is defined as an aggregated composite construct of several dimensional constructs measured by Financial DEA. When the Financial DEA constructs do not fully cover the domain of the firm performance phenomenon, disturbances are introduced to the domain of firm performance and therefore alter the meaning. In Financial DEA, disturbances may be due to the difference between the domain of the firm performance phenomenon and the composite domain formed by dimensional constructs. However, the impact on Financial DEA research is different due to different perspectives taken. Through a formative lens, the interpretation of firm performance is no longer operational efficiency but changes to financial efficiency. The remedy is to ensure that the dimensional constructs measured share the same domain as the firm performance phenomenon intended to measure. For example, refining dimensional construct from financial efficiency to operational efficiency by deleting the attributes related

to activities other than operational efficiency can move the domain measured closer to the operational efficiency, which is the domain of the firm performance phenomenon intended to measure from financial efficiency, which is the domain actually measured.

In a reflective model of the relationship between variables and construct in the operationalisation phase (Quadrant I and II in Figure 6 - 4), the dimensional construct reflects the physical production process similar to conventional DEA. The accounting variables in the Financial DEA models introduce random measurement errors due to the differences between physical measures and accounting measures. For example, in the continuum in Figure 6 - 3, from the reflective end to the formative end, heterogeneity is introduced by various sources such as prices, accounting choices, and business activities. This heterogeneity is a methodological issue in DEA (Dyson et al., 2001). The heterogeneity introduced along the continuum could influence how well the physical process is measured. The interpretation of the Financial DEA result is around the physical productive efficiency. However, the measurement errors could lead to numerically biased quantitative results. This numerical bias is due to the difference between physical measures and accounting measures. The remedy is to use alternative accounting measures to proxy the same physical measures in separate Financial DEA models. The covariance between these Financial DEA results can effectively reduce the measurement errors. For example, to measure operational efficiency, alternative accounting measures including gross property plant and equipment, net property plant and equipment, and depreciation expenses can be used to calculate the construct. The covariance of the constructs represents the construct of operational efficiency and reduces the impact of accounting choices.

In a formative model of the relationship between variables and construct in the operationalisation phase (Quadrant III and IV in Figure 6 - 4), the dimensional constructs are the composite constructs defined by the indicators represented by the accounting variables. The combination of accounting variables introduces disturbances into the composite dimensional construct. A disturbance is a difference between the domain of the dimensional construct that is supposed to be measured and what is actually measured. Variable selection has long been a methodological issue in DEA (Dyson et al., 2001). In the context of Financial DEA, variable selection is still a critical issue (section 4.3.4). For example, to measure the firm performance of a capital-intensive industry, the attribute of capital needs to be in the Financial DEA model. Omitting such an attribute would alter the domain being measured.

The differences between the domain supposed to be measured and the domain that Financial DEA actually measures are the source of disturbances. The disturbances lead to the misinterpretation of Financial DEA results due to changes in the domain of the dimensional construct. The remedy is to select accounting variables based on the attributes decided by the domain that researchers aim to measure. In Financial DEA, accounting variables need to cover the key business activities and the industry features to measure financial efficiency at the firm level. For instance, a labour-intensive industry needs to cover the measure of labour resources at the firm level. Within the firm, both primary and support activities need to be considered. For instance, not only does key operational efficiency need to be covered, the support activity such as technology development activities also need to be covered.

6.3.4. Summary

This section answers *RQ2* by synthesising the methodological issues (section 4.3), the dimensional constructs (section 4.2.1), with the measurement models (section 2.5) to provide a conceptual framework for the Financial DEA modelling process. First, based on the measurement models in section 2.5, a continuum between the reflective and formative constructs in the context of business structure is added (Figure 6 - 3) to the framework of dimensional constructs discussed in *RQ1*. Second, the measurement models and Financial DEA studies were synthesised into a four-quadrant framework (Figure 6 - 4) to classify the Financial DEA models into four main categories. These four quadrants of Financial DEA modelling also identify the sources of measurement errors with alternative modelling lenses. The next section continues the discussion of measurement errors and provides empirical illustrations that quantify the impact on the results of Financial DEA.

6.4. Research Question Three

This section answers the third research question:

> ***RQ3.*** *What are the empirical impacts of methodological choices on the results of Financial DEA?*

This section synthesises the findings of the three empirical tests discussed in Chapter 5 with the conceptual framework of Financial DEA at the construct level in *RQ1* and the four-quadrant framework with measurement error sources (conceptual framework of Financial

DEA at the modelling level) in *RQ2*. This section further discusses the empirical impacts of the methodological issues on the results of Financial DEA considering various interrelationships between dimensional constructs in a business structure.

Overall, as illustrated in Figure 6 - 5, four categories of factors are found at various levels within the context of a business structure. The first is that the price factors locate at the operational level when expanding from the physical production level. The second is that accounting choices are located at the activity level when expanding from the operational level. Third, the operational characteristics locate at the firm level when expanding from the activity level. Finally, the industry characteristics locate at the market level when expanding from the firm level.

This section is structured as follows. Section 6.4.1 discusses the empirical impact of variations in the price factor on the results of Financial DEA. Section 6.4.2 discusses the empirical impact of accounting choices on the results of Financial DEA. Section 6.4.3 discusses the empirical impact of operational characteristics on the results of Financial DEA. Section 6.4.4 discusses the empirical impact of industry characteristics on the results of Financial DEA. Section 6.4.5 summarises this section.

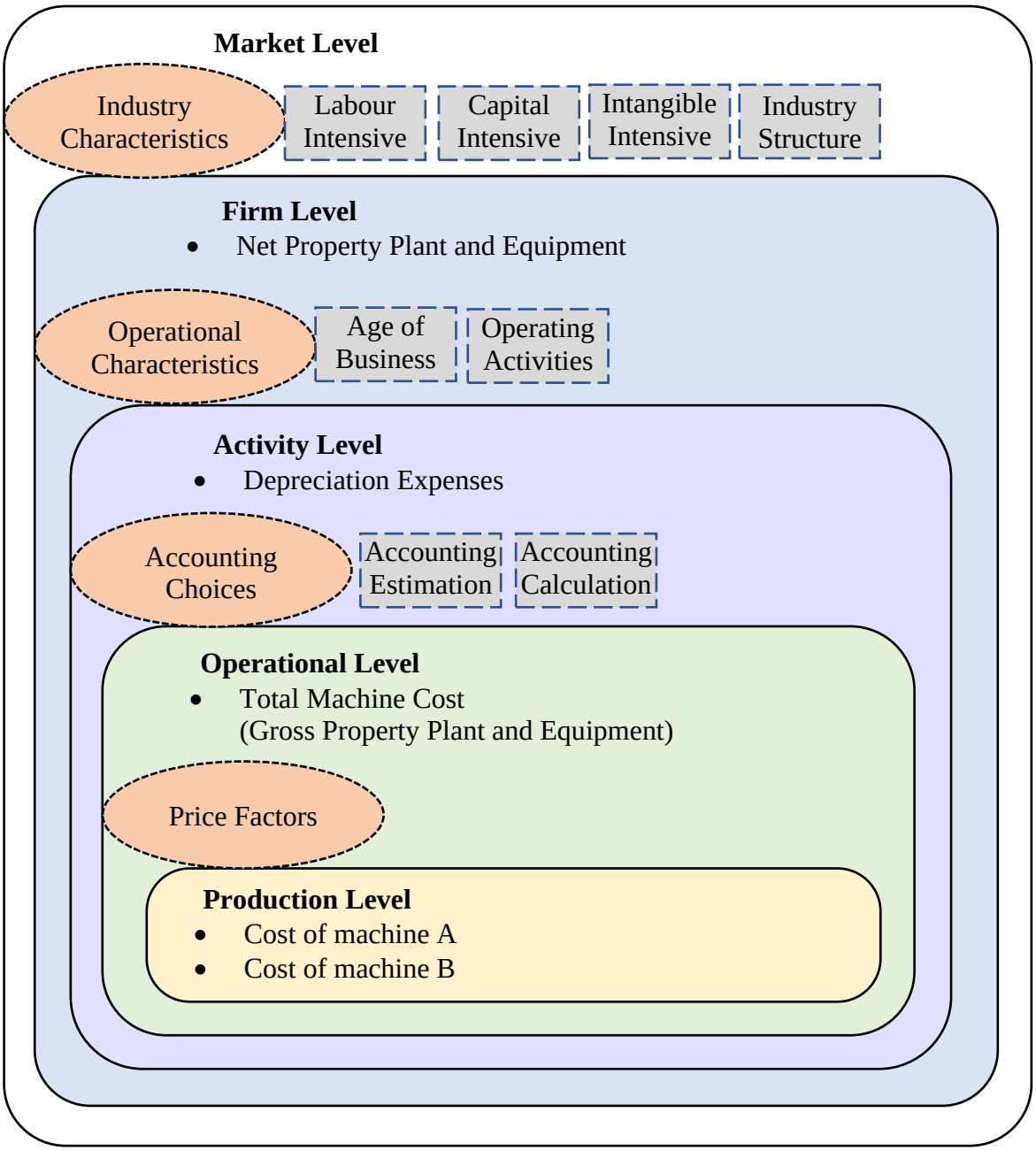


Figure 6 - 5 Conceptual Framework of Financial DEA at the Factor Level

6.4.1. Price Factors at the Operational Level

The inclusion of the price factors expands the factors of production from being measured by physical measures in conventional DEA to being measured by price-based aggregated accounting variables. The literature reviewed in Chapter 2 (section 2.2.2.2) discussed conceptually how the price-based aggregated accounting data could bias the results of technical efficiency (Farrell, 1957; Zelenyuk, 2020). The test in section 5.2 empirically examines the magnitude of price variation on the Financial DEA models with simulated data.

The study used simulated data to empirically quantify the impact of price-based aggregated accounting data on Financial DEA when measuring technical efficiency. The basis for the test is that accounting variables are the product of physical measures and prices. As illustrated in Figure 6 - 5, price-based aggregated accounting variables include price factors in the performance measurement. The inclusion of price factors expands the domain from the single physical production process (Figure 6 - 5, the production level, yellow box) to the operational level where production processes are aggregated (Figure 6 - 5, the operational level green box). For example, the total machine cost, gross property, and plant and equipment (accounting variable) represent the number of machines (physical measure) multiplied by the cost of each machine (prices).

The empirical test in section 5.2 was designed using the lens of reflective modelling. The test used a simulated Monte Carlo experiment to empirically quantify the potential measurement errors due to price variation when using Financial DEA models to estimate technical efficiency.

With the reflective lens, the Financial DEA models closely reflect underlying physical production processes (yellow box, Figure 6 - 5). The reflective view is in line with conventional DEA in that the model aims to measure the technical efficiency of a production process using only physical measures (Farrell, 1957). Compared with conventional DEA in the yellow box in Figure 6 - 5, the productive efficiency in the Financial DEA literature is essentially technical efficiency as defined by Farrell (1957), which is the “success in producing as large as possible an output from a given set of inputs” (Farrell, 1957, p. 254). According to Farrell (1957), three types of efficiencies exist for the efficiency of the production process: technical efficiency, allocative efficiency, and overall efficiency. Technical efficiency is distinguished from the other two efficiencies in that “price (allocative)

efficiency measures the extent of a firm's adaptation to a particular set of prices" (Farrell, 1957, p. 261). However, "price (allocative) efficiency is a measure that is both unstable and dubious of interpretation; its virtue lies in leaving technical efficiency free of these faults, rather than in any intrinsic usefulness" (Farrell, 1957, p. 261).

Section 5.2 illustrates that with incremental variations in price factors, the correlation between the results of the Financial DEA and the physical technical efficiency decreases. The price factor represents random measurement errors introduced to the variables due to the difference between accounting and physical measurements.

The test finds that different price variations impact the results of Financial DEA to different degrees. When the price factor is constant, the results of the Financial DEA is the same as the technical efficiency. However, when prices vary from a narrow to a broad degree, the impact on the results of Financial DEA increases. When the sample size is relatively large, and the price factor is relatively homogeneous (i.e. the variation of the price factor is narrow), the Financial DEA estimates and the technical efficiency are significantly correlated with high coefficients. However, when the sample size is small, or the prices vary widely (i.e. the variation of the price factor is broad), Financial DEA does not provide an accurate estimation of the underlying technical efficiency.

When accounting variables proxy the production elements, the variation in price functions for individual DMUs will impact the results of Financial DEA models. Researchers need to examine price heterogeneity when constructing samples. For instance, firms with low buying and selling powers face similar price functions for inputs and outputs. The Financial DEA estimates are likely to reflect underlying technical efficiency closely. By comparison, firms that have high buying and selling powers are likely to face wide price variations. Financial DEA will provide a much less accurate estimation of the underlying technical efficiency. When researchers use accounting variables under these conditions, firms need to be classified into sub-samples with less price heterogeneity.

The test empirically illustrates the change of results when using price-based aggregated accounting variables in Financial DEA to estimate technical efficiency. Researchers have discussed conceptually the Financial DEA estimation and how it can bias technical efficiency (Färe & Zelenyuk, 2002; Färe et al., 2004; Zelenyuk, 2020). This test provides additional empirical evidence to quantify the change. The finding can guide future researchers who wish

to use price-based aggregated accounting data to estimate the underlying technical efficiency without significant biases. For example, researchers can sort the sample into subsets with similar prices or pricing strategies when a sample set is mixed with a broad variation in prices. This study also provides empirical evidence to evaluate the literature that has stated that aggregated accounting data does not bias the estimation of technical efficiency (Banker et al., 2007). Only when certain conditions are met, such as when buying and selling powers are relatively weak and the market is relatively large, Financial DEA may provide a reasonable estimation of technical efficiency.

The test in section 5.2 may provide additional insight into benchmarking firm performance with the lens of formative modelling. With the formative modelling lens, the Financial DEA models are assumed to form a composite construct. Formative models may introduce disturbances to the composite construct when the domain defined by the indicators is different from the construct to be measured. The test in section 5.2 can be viewed as forming a composite construct comprising a mix of physical and additional price information for managerial decision making. The variation of prices introduces additional price information into the DEA results. The price information is regarded as an additional attribute and diverges the Financial DEA estimated results from using the conventional DEA model with physical measures.

In summary, from the lens of reflective modelling, the price-based aggregated accounting variables in Financial DEA introduce measurement errors in estimating the underlying technical efficiency. However, from the lens of formative modelling, the aggregated accounting variable incorporates the price information and physical information into the composite construct. The price-based aggregated accounting variables expand the meaning of the construct from technical efficiency to include various price information.

6.4.2. Accounting Choices at the Activity Level

Accounting choices introduce additional information such as accounting estimation and accounting calculation into the dimensional constructs at the activity level measured by Financial DEA. In Figure 6 - 5, the inclusion of accounting choices expands the domain from the operational level (green box) to the activity level (purple box). Section 5.3 summarised accounting variables that measure the amount and use of capital and examined the impact of heterogeneity of accounting choices on Financial DEA. The test in section 5.3 starts from the

operational level (green box), where gross property plant and equipment (*GPPE*) measures the capital, including price factors. Compared with *GPPE*, depreciation expenses *(DP)* are also impacted by accounting choices, including accounting estimation and the accounting calculation method. The accounting estimation of useful life *(UL)* depends on the estimated utilisation rate of a machine. The depreciation calculation method depends on the accounting choice to best describe the pattern of activities.

With the lens of reflective modelling, the Financial DEA models in section 5.3 are assumed to reflect underlying physical production processes closely. Compared with *GPPE*, *DP* introduces information on accounting estimation and the calculation method. This information is regarded as measurement error and leads to divergence of the Financial DEA results from the results where only the price factor is covered (*GPPE*, the green box in Figure 6 - 5). As evident by section 5.3.1, the wider the variation of the accounting choices, the further the Financial DEA results measured with *DP* diverged from the results measured with *GPPE*. When researchers aim to calculate the efficiency of a physical production process (the yellow box), the information related to accounting choices in *DP* may not necessarily be useful. By comparison, *GPPE* is only impacted by the price factors. The heterogeneity of accounting choices is regarded as the source of random measurement errors. In a reflective model, the information of accounting choices may lead to the difference between accounting and physical measures.

With the lens of formative modelling, the Financial DEA models in section 5.3 are assumed to form a composite construct to benchmark firm performance. Compared with *GPPE*, *DP* introduces information on accounting estimation and the calculation method. This information is regarded as disturbances and leads to a different domain of the construct. Compared with using *GPPE*, when using *DP*, the domain expanded from the operational level (green box) to the activity level (purple box) in Figure 6 - 5. As evident by section 5.3.2, in the gold industry, the estimated useful life varied widely due to the diversity of the ore sites. Also, the method of depreciation is unit-of-activity, characterised by variation in the amount due to the significant productivity variation from one year to another (Weygandt et al., 2015). This information was incorporated into the Financial DEA, and Financial DEA results measured by *DP* diverged from the results measured by *GPPE*. The information on accounting choices is the key to distinguishing the domain that is aimed to be measured and the domain that is practically measured by Financial DEA. In section 5.3, accounting choices

added attributes to expand the composite construct from only including price factors to including the accounting information (the green box to the purple box, Figure 6 - 5). The expansion of the domain can provide more informative results for managerial decision making.

In sum, from the view of reflective modelling, when researchers aim to calculate the efficiency of the underlying physical production process, the accounting choices could introduce measurement errors to the physical measurements and lead to divergence between the Financial DEA results and the results obtained using physical measures of productivity. In a reflective model, the underlying construct is physical productive efficiency, represented by indicators of production elements, such as capital, labour, and materials. In conventional DEA, these indicators are usually represented by physical variables and quantified by physical measures. However, in Financial DEA, indicators are represented by accounting variables and quantified by accounting measures. In Financial DEA, the “true” measure is decided by the underlying physical production process, and variations are due to differences between physical and alternative accounting measures.

From the view of formative modelling, researchers may need to calculate a composite construct of productive efficiency with additional accounting choices for particular managerial decision-making purposes. The additional accounting choices introduced by alternative accounting variables may provide useful information, leading to results that are different from physical productivity. Researchers need to select the accounting variable that suits the research purpose. From the view of formative modelling, this test may provide additional insight into the impact of accounting variable choice on benchmarked performance. Under formative modelling, the Financial DEA models are assumed to form a composite construct, a mix of physical productivity with additional accounting choice information needed for managerial decision making. The choice of stock and flow accounting variables introduces various accounting and operational factors into Financial DEA results. They are regarded as additional attributes of the composite construct and lead to divergence of the Financial DEA measured results from the results obtained using a physical productivity DEA model.

6.4.3. Operational Characteristics at the Firm Level

The operational characteristics introduce additional information such as operating activities and business ages into the construct of Financial DEA. In Figure 6 - 5, the inclusion of operational characteristics expands from the activity level (purple box) to the firm level (blue box). The accounting variable, net property plant and equipment (*NPPE*), as a measure of capital, empirically examined the impact of the heterogeneity of operational characteristics on Financial DEA. The test in section 5.3 starts from the operational level (green box), where *GPPE* measures the capital with price factors. After comparing *DP* with *GPPE* (section 6.4.2), *NPPE* was also compared. *NPPE* is impacted by accounting choices and operational characteristics, such as the age of business and the operating activities.

With the lens of reflective modelling, the Financial DEA models in section 5.3.1 are assumed to reflect underlying physical production processes closely. Compared with *GPPE*, *NPPE* introduces information on accounting choices and operational characteristics. This information is regarded as measurement errors and leads to divergence of the Financial DEA results from the results where only the price factor is incorporated (*GPPE*, the green box in Figure 6 - 5). As evident in section 5.3.1, the wider the variation of the accounting choices and operational characteristics, the further the Financial DEA results measured with *NPPE* diverged from the results measured with *GPPE*. Compared with *DP* (section 6.4.2), *NPPE* generated results that diverged more from *GPPE* than *DP*. When researchers aim to calculate the efficiency of the physical production process (the yellow box in Figure 6 - 5), the information on accounting choices and operational characteristics in *NPPE* is not necessarily helpful. By comparison, *GPPE* may provide a closer estimation of the physical technical efficiency. The heterogeneity of accounting choices and operational characteristics is regarded as the source of random measurement errors. In a reflective model, operational characteristics may lead to differences between accounting and physical measures.

When synthesising with the discussion from section 6.4.2, *GPPE* and *DP* produce similar Financial DEA results, and *NPPE* introduces the greatest variation (section 5.3.1). Accounting choices impact *DP*, and both accounting choices and operational information impact *NPPE*. When researchers aim to calculate the technical efficiency of the underlying physical production process, the operational information and accounting information in *DP* and *NPPE* is not necessarily useful in measuring the physical productive efficiency. Instead,

the accounting information and operational information are regarded as the source of random measurement errors.

With the lens of formative modelling, the Financial DEA models in section 5.3 are assumed to form a composite construct to benchmark firm performance. Compared with *GPPE*, *NPPE* introduces information on both accounting choices and operational characteristics. This information is treated as a disturbance in the composite construct and alters the domain measured. Compared with *GPPE*, when using *NPPE*, the domain expanded from the operational level (green box in Figure 6 - 5) to the firm level (blue box in Figure 6 - 5). This finding is evident in section 5.3.2, in the gold industry, where the business age varied widely. Also, mining activities were difficult to compare due to the unique nature of ore sites. This information was incorporated into the Financial DEA, and the results measured by *NPPE* diverged from the results measured by *GPPE*. The information of accounting choices and operational characteristics distinguish the domain aimed to measure and the domain that Financial DEA actually measures. In section 5.3.1, the operational characteristics added attributes in addition to *DP* and expanded the composite construct from only including price factors (*GPPE*) to also including accounting information (*DP*) and operational characteristics (*NPPE*). The expansion of the domain can provide more informative results for managerial decision making.

When synthesising with the discussion from section 6.4.2, *NPPE* is the more informative accounting variable than *DP*, followed by *GPPE*. *NPPE* is a function of relative age and contains information on both *AGE* and *UL*. *NPPE* represents the remaining value of a machine. When researchers are interested in future capacity, *NPPE* provides relevant information for this research purpose. *DP* is a function of *UL* and contains the accounting information of estimated *UL*. When researchers are interested in the utilisation rate of a machine a year, *DP* is a more suitable measure.

The test in section 5.3 demonstrates that accounting choices (section 6.4.2) and operational information (section 6.4.3) introduce additional information to the physical measures used in Financial DEA and influence the results. To measure the capital in a production process, *GPPE* is arguably the most stable measure, given it is influenced the least by accounting choices and operational factors. However, *GPPE* does not contain information on the utilisation rate, neither at the activity level nor at the firm level (Figure 6 - 5). In industries where plant usage fluctuates from year to year. For example, in the gold industry, *GPPE*

cannot proxy for utilisation information accurately. *NPPE* incorporates operational information at the firm level, utilisation information at the activity level, and historical cost at the operational level. However, *NPPE* represents the value in use for the future. For industries where the plant is relatively new or where the non-depreciable portion is high, for example, in the gold industry, the *NPPE* may generate very similar results to *GPPE* in Financial DEA. *DP* captures the usage of capital over a year at the activity level. For example, in the gold industry, where the depreciation method is the units-of-activity, capital utilisation varies significantly from one year to another and is not a fixed proportion of the cost (i.e. straight-line depreciation method). Hence, *DP* can generate very different Financial DEA results compared to *GPPE* and *NPPE*.

In sum, from the view of reflective modelling, when researchers aim to calculate the efficiency of the underlying physical production process, accounting choices (section 6.4.2) and operational information (section 6.4.3) could introduce measurement errors to the physical measure. These errors could lead to divergence between the Financial DEA results and the results obtained using physical measures. In a reflective model, the underlying construct is physical productive efficiency, represented by indicators of production elements, such as capital, labour, and materials. In conventional DEA, these indicators are usually represented by physical variables and quantified by physical measures. However, in Financial DEA, indicators are represented by accounting variables and quantified by accounting measures. In Financial DEA, the "true" measure is decided by the underlying physical production process, and variations are due to the differences at the activity level (purple box in Figure 6 - 5) and the firm level (blue box in Figure 6 - 5) between physical and alternative forms of accounting measures.

From the view of formative modelling, researchers may need to calculate a composite construct of productive efficiency with additional accounting choices (purple box in Figure 6 - 5) and operational information (blue box in Figure 6 - 5) for particular managerial decision-making purposes. The additional accounting choices and operational factors introduced by alternative accounting variables may provide useful information, leading to different results from physical productivity. Researchers need to select the accounting variable that suits the research purpose. Using a formative lens, this test may provide additional insight into the impact of accounting variable choices on benchmarked performance. Under formative modelling, the Financial DEA models are assumed to form a composite construct, a mix of

physical productivity with additional accounting (purple box in Figure 6 - 5) and operational information (blue box in Figure 6 - 5) needed for managerial decision making. The choice of stock and flow accounting variables introduces various accounting and operational factors into Financial DEA results, which are regarded as additional attributes of the composite construct and lead to divergence of the Financial DEA results measured from the results obtained using a physical productivity DEA model.

6.4.4. Industry Characteristics at the Market Level

The industry characteristics introduce additional information such as the intensity of resources and the industry structure into the construct of Financial DEA at the market level. Figure 6 - 5 shows that the inclusion of industry characteristics expands from the firm level (blue box) to the market level (white box). The test in section 5.4 examined the impact of alternative accounting variables on the results of Financial DEA in six industries with unique features. The results found that the industry characteristics impacted the comparability of the Financial DEA results. The more homogeneous the firms were within an industry, the stronger the convergent validity identified was in the Financial DEA results. The more heterogeneous the firms were within an industry, the stronger the discriminant validity was across the different Financial DEA results.

With the lens of reflective modelling, at the industry level, the accounting variables only proxy the physical production process indirectly. In section 5.4, the variables were altered when testing the convergent validity, but the indicators remained the same. It can be argued that the constructs measured in Test A reflect the same underlying construct. However, due to the individual industry features, the results varied. For example, the omission of intangible assets impacted the food industry and the personal services industry relatively more than the other industries. These two industries have a relatively high intensity of intangible assets. The information at the industry level can diverge the Financial DEA results further from the physical production process (yellow box in Figure 6 - 5). In a reflective model, the information on industry characteristics is regarded as random measurement errors that lead to differences between accounting and physical measures.

With the lens of formative modelling, Financial DEA is used to measure firm performance as a benchmarking tool (section 2.2.3). At the industry level, the relationship between the accounting variables and the physical production elements is indirect. Rather, the

benchmarking feature of Financial DEA is emphasised. The inputs are the attributes to minimise, representing the less-the-better performance measures. The outputs are the attributes to maximise, representing the more-the-better performance measures (Cook et al., 2014).

The test in section 5.4.1.3 focused on construct validity and examined the impact of different degrees of disturbances due to alternative accounting variables on the construct validity of the Financial DEA results. The test found that industry characteristics impacted the magnitude of discriminant validity of Financial DEA results. For example, the relatively more heterogeneous industries, such as the clothing, food, and gold industries, showed stronger discriminant validity among alternative variable selection than the other industries. Financial DEA captured the unique industry features at the industry level. By expanding the domain from the firm level (blue box in Figure 6 - 5) to the industry level (white box in Figure 6 - 5), industry characteristics provided additional attributes to the composite construct. As a result, the new construct is more informative and may provide more comprehensive managerial decision-making information.

The test in section 5.4 examines the impact of industry characteristics on the construct validity of Financial DEA models. This test found that the research context influences convergent validity and discriminant validity across different Financial DEA models. Models constructed using similar indicators and measures had the highest convergent validity levels among the Financial DEA models. Additionally, the more homogeneous the samples were, the stronger the convergent validity identified in the Financial DEA results. Furthermore, the more heterogeneous the sample sets were, the stronger the discriminant validity identified across the different Financial DEA results.

6.4.5. Summary

Section 6.4 answers *RQ3* by synthesising the empirical findings of the three tests in Chapter 5 with the conceptual framework of Financial DEA at the construct level developed in *RQ1* and the four-quadrant frameworks with the measurement errors developed in *RQ2*. This section discusses selective empirical impacts of methodological issues on the results of Financial DEA in the context of a business structure. Overall, four categories of factors that could impact the Financial DEA results are found at five levels in a business structure, from the

production level to the market level (Figure 6 - 5). The empirical impact of each factor is discussed with alternative modelling lenses.

6.5. Chapter Summary

This chapter discussed the findings of this study and synthesised them into a comprehensive theoretical foundation for Financial DEA. This theoretical foundation synthesised the dimensional constructs measured by Financial DEA into a business structure and identified the sources of methodological issues in the application. Selective methodological issues were further examined in the context of a business structure.

First, *RQ1* (section 6.2) provided a conceptual framework at the construct level, locating the dimensional constructs in section 4.2.1 into the business structure. The dimensional constructs were able to be mapped against Porter's value chain framework (section 2.4.2), and these constructs identified the development trend of Financial DEA that mainly flourished in the research areas at the operational and firm levels. Second, *RQ2* (section 6.3) provided a conceptual framework at the modelling level, which identified four quadrants with the reflective/formative modelling lenses. With alternative lenses, the measurement errors could be interpreted differently. Third, *RQ3* (section 6.4) provided a conceptual framework at the factor level, which identified four types of factors at five aggregation levels in the firm structure. Three sets of tests illustrated the empirical impacts of these factors on Financial DEA. The conceptual framework at the factor level classified the source of errors according to various levels of firm structure. With alternative modelling lenses, different stories can be told. These three conceptual frameworks build the conceptual foundation for Financial DEA, which can guide Financial DEA applications.

Chapter 7: Conclusion

7.1. Chapter Introduction

This study has developed a conceptual foundation for Financial DEA. The study was motivated by the need to understand how Financial DEA can be used for firm performance measurement in various business contexts. Combining accounting information into DEA brings challenges and opportunities to the emerging Financial DEA research. This study provides a comprehensive conceptual foundation for Financial DEA by structuring the Financial DEA literature at the construct level, classifying the Financial DEA models into four quadrants at the modelling level, and examining the empirical impact of measurement errors at the factor level.

Chapter 7 is structured as follows. Section 7.2 reviews the study, including the methods and contributions. Section 7.3 discusses the limitations and future research. Section 7.4 concludes.

7.2. Summary of the Study

7.2.1. Methods

To develop a conceptual foundation of Financial DEA, this study applied a two-phase research design. In the first phase, a typology, a construct level conceptual framework, and a modelling level conceptual framework were developed from the literature. For the Financial DEA typology, 248 models were analysed from 1990 to 2021. The Financial DEA models were examined to understand the diversity of Financial DEA models from the first to the most recent models. The findings were synthesised into a literature typology, which comprehensively described the variety of Financial DEA applications at the dimensional construct level. The typology identified 12 dimensional constructs and nine indicators representing the existence, or absence, of dimensional constructs.

For the conceptual framework at the modelling level, the literature of the measurement models was adapted from sociology and psychology. The measurement models discussed the relationship between concepts, constructs, indicators, and variables, which serve the goals of

modelling Financial DEA in the conceptual framework. Findings from the literature were synthesised into a conceptual framework at the modelling level, which described the methodological issues in Financial DEA. The conceptual framework found four quadrants of measurement models of Financial DEA. In each quadrant, measurement errors were highlighted. These measurement errors were illustrated empirically in phase two selectively.

In the second phase, three selected methodological issues identified in the first phase were examined using simulated and archival data. The simulated data were used to isolate the relationships between factors of interest. The archival data were used to ensure the tests represented reality. The samples were designed to capture the various key factors in each test with various sample sizes. The archival data also covered a variety of business features with a range of levels of homogeneity. Results were consistent with the propositions that the accounting variables, compared with the physical variables in conventional DEA, are impacted by various factors in the context of a business structure.

7.2.2. Main Outputs and Contributions

The study provides three main outputs, which matched the three research questions. First, a conceptual framework at the construct level was developed. By reviewing the literature of Financial DEA, a comprehensive literature typology was developed, which forms a conceptual framework for Financial DEA at the construct level, describing the constructs measured in the context of a business structure. Second, a conceptual framework at the modelling level was developed. A four-quadrant conceptual framework was developed by organising the measurement models and measurement errors, thereby locating the measurement errors in Financial DEA studies. Third, a conceptual framework at the variable level was provided. By conducting a series of empirical tests, measurement errors under alternative modelling lenses were examined. This section synthesises the outputs and the contributions.

7.2.2.1. RQ1 – conceptual framework at the construct level

The first output is a conceptual framework at the construct level that classifies the dimensional constructs of firm performance measured by Financial DEA. As a result, 12 dimensions of firm performance were identified in the Financial DEA literature. These dimensional constructs can be structured in a business context. There were two dimensional constructs at the production level, being productive efficiency and intermediation efficiency.

There was one dimensional construct at the operational level, being operational efficiency. There were six dimensional constructs at the activity level: marketing efficiency, compensation efficiency, funding efficiency, asset efficiency, total asset efficiency, and current asset efficiency. There were two dimensional constructs at the firm level, being financial efficiency and profitability efficiency. There was one dimensional construct at the market level, being share-market efficiency. To represent the existence or absence of the 12 dimensional constructs, nine indicators were found, being labour, material, capital, assets, liabilities, equity, expenses, production volume, and revenue.

This output structures the Financial DEA literature within a business structure that provides an overview of the scope of Financial DEA to date. This framework is a prerequisite for the sources of methodological issues in Financial DEA.

The findings contribute to the literature by facilitating the adaptation of Financial DEA models to firm performance measurement and financial statement analysis. Researchers have been using Financial DEA for these purposes for the past thirty years with broad implications of various perspectives of firm performance (Demerjian et al., 2013; Smith, 1990). However, the Financial DEA studies are not underpinned by a clearly articulated conceptual basis, and therefore, an overview of the development of Financial DEA is needed.

This framework contributes to the literature by structuring the Financial DEA studies and providing an overview of Financial DEA application in a business context. The framework expands the dimensions measured by Financial DEA from the physical production level in the conventional DEA to the firm level and market level. The framework horizontally covers various industries, ranging from financial to general industries. The framework vertically covers a range of business structures, including specific production lines and also across industries. This framework at the construct level describes the development of Financial DEA applications to-date that researchers can use as a map for structuring Financial DEA studies.

7.2.2.2. RQ2 – conceptual framework at the modelling level

The second output of this study is a conceptual framework at the modelling level that identifies methodological issues in Financial DEA applications. Four quadrants of measurement models were identified within this conceptual framework. Each quadrant was formed by either the reflective or formative models in explaining the relationships between variables, constructs, and phenomena.

In a reflective model, the measurement errors are random measurement errors. In the relationship between constructs and phenomena, random measurement errors affect the dimensional constructs. This mismatch might bias the research findings if the dimensional construct is not the only dimension and so could not perfectly encompass the meaning of firm performance. In the relationship between variables and constructs, random measurement errors affect the variables. Factors such as prices and accounting choices may cause the accounting measures to diverge from physical measures.

In a formative model, the measurement errors are disturbances. In the relationship between constructs and phenomena, disturbances affect the domain of the phenomenon of firm performance. When the dimensional constructs cannot cover the domain of the phenomena, the meaning of the phenomena may alter. In the relationship between variables and constructs, disturbances affect the domain of constructs. For example, if indicators were operationalised into accounting variables with errors, the meaning of the composite construct could be biased.

This study connects the Financial DEA modelling with measurement models using two lenses of modelling. The study accelerates the understanding of Financial DEA modelling as a firm performance measurement tool. However, by nature, different modelling lenses lead to different types of measurement errors and interpretation consequences. As a result, misspecification of the Financial DEA model may affect theory development and prohibit researchers from meaningfully testing theory (Edwards & Bagozzi, 2000).

The conceptual framework at the modelling level contributes to the Financial DEA literature by providing a broad classification of measurement models and measurement errors. The conceptual framework can be used to appraise the emerging Financial DEA research. The conceptual framework can also be used to design future Financial DEA research to minimise measurement errors. The Financial DEA research can be classified into specific quadrants

depending on the research goal and research design. For each quadrant, the methodological issues are highlighted at various locations of measurement models. Researchers can use the conceptual framework to diagnose the potential methodological issues in an empirical Financial DEA study. The framework facilitates the development of Financial DEA by highlighting the pitfalls in the application procedure.

7.2.2.3. RQ3 – conceptual framework at the factor level

The third output of the study is a conceptual framework at the factor level. A series of quantitative tests demonstrated the empirical impact of selected methodological issues on Financial DEA results. Three sets of empirical tests with simulated and archival data were conducted. The selected methodological issues were quantified in the relationship between variables and construct since the relationship between constructs and phenomena were theoretical and not suitable for quantification. The methodological issues examined by the three tests were the impacts of (a) the price factors, (b) the forms of accounting variables, and (c) alternative accounting variables in models on the results of Financial DEA across various industry settings.

First, the price factor reduced the accuracy of Financial DEA from the view of reflective modelling. Compared with conventional DEA, Financial DEA generated a less accurate estimation of the physical technical efficiency. The estimation was relatively inaccurate when price factors varied to a wider degree in a market with high selling and buying powers. However, from a formative modelling perspective, the price factor captured more information in the accounting measures in Financial DEA than the physical measures in conventional DEA.

Second, the choice of stock and flow forms of accounting variables led to changes in the Financial DEA results. From the view of reflective modelling, various forms of accounting variables are essentially due to factors such as accounting information and operational characteristics in accounting measures compared to physical measures. The greater the degree of the variation in the factors, or the more factors influencing the accounting measures, the further Financial DEA diverged from conventional DEA. However, from a formative modelling perspective, various accounting and operational factors delivered additional information to Financial DEA rather than just the physical measures used in conventional

DEA. The estimated efficiency scores were the composite of physical technical efficiency with varied accounting information and operational characteristics.

Third, the choice of alternative accounting variables in the various industry features impacted the construct validity of Financial DEA in the formative modelling. With the change of the accounting variables, the construct validity was impacted. The specification of Financial DEA models, which include the indicators and the accounting variables, needs to align with the targeted composite construct of the research goal.

The findings also provide guidance and protocols to researchers for practical Financial DEA applications when applying Financial DEA models. Researchers suggest that DEA models need to be applied with care due to various method and methodological issues (Dyson et al., 2001). Researchers have also suggested that DEA models have different features when building a productive efficiency frontier or benchmarking frontier (Cook et al., 2014). This study uses the literature of DEA methodology and measurement models (section 4.3) to investigate the methodological issues and protocols when applying Financial DEA. The findings emphasise that four types of factors in the business structure are the key to ensure meaningful financial DEA results. The four types of factors are (a) price factors, (b) accounting choices, (c) operational characteristics, and (d) industry characteristics.

The findings contribute to the Financial DEA literature by providing empirical evidence on the magnitude of key methodological issues of the four types of factors in the conceptual framework. The measurement errors in reflective and formative Financial DEA models are of different natures, impacting the results differently. Building a Financial DEA model without being aware of the pitfalls may affect theories being meaningfully tested.

7.3. Limitations and Future Research

7.3.1. Research Cycle

Under the positivist view, research has a cyclical feature through inductive and deductive research that starts with facts of one cycle and ends with facts beginning the next cycle. The first step – induction, draws a general conclusion from empirical observations. The second step – deduction, is a logical analysis of general theory through empirical scrutiny (Arbnor & Bjerke, 2009). However, this study only selectively covered the deductive research phase due

to its exploratory nature. Based on the conceptual framework at the modelling level, generated by *RQ2*, selective but not exhaustive measurement errors were quantified by empirical tests in *RQ3*.

Future research studies can be done by carrying out additional deductive research to provide further structured explanations. Based on the conceptual framework, research questions, propositions, and hypotheses can be built and operationalised to test. Empirical tests are advocated by empiricists who believe that knowledge is based on experience (either qualitative or quantitative data), as opposed to base on pure logical relations (rationalism) (Goodwin & Goodwin, 2017; Ryan et al., 2002). For each conceptual quadrant, measurement errors exist at various locations, depending on the type of measurement model. The measurement errors can be specified into a range of factors that impact the Financial DEA results. For example, price is one factor that could impact the Financial DEA results in the first-order reflective modelling. A comprehensive list of factors could be explored, and a range of tests carried out to explain the impact of factors on Financial DEA results.

7.3.2. Quantitative Data

The quantitative data used in the empirical tests were simulated data and archival data. To generate simulated data, a range of assumptions need to be made for the parametric models and the simulated data distribution. However, the predetermined parameters may not cover all possible factors. The scope of simulated tests can only cover and isolate the most influential factors in a specific research context. Furthermore, the relationship between factors can be too complicated to simulate. For instance, in the test of alternative stock and flow forms of accounting variables (section 5.3), the two most influential factors were assessed age (the operational characteristics) and the useful life (the accounting information). However, the two factors were related since, for one machine, its age could not exceed the useful life. Other factors also impact Financial DEA results in addition to the two most influential factors simulated in the test. As identified in the empirical tests, industry features such as the specific depreciation methods also impact the Financial DEA results generated by alternative forms of accounting variables. Due to the design of the simulated test, only one type of accounting depreciation method was assumed for the comparability of the results. The most common depreciation method, the straight-line method, was chosen.

The study also conducted empirical tests with archival data. Due to the assumptions and parametric functions, simulated data incorporate weakness in that they may not represent reality. While the archival data remedy the weakness of simulated data, they are exposed to weakness too. The archival data only covered selective industries and a five-year period, limiting the generalisation of the findings. The industries were selected to be relatively homogenous (i.e. the average efficiency scores are above 0.7 in Demerjian (2018)) with relatively different homogeneity levels within the selected industries and various industry features. Due to the algorithm features of DEA, extremely heterogenous data can bias the results. To ensure the reasonableness of the results, only relatively homogenous industries were selected. To test the impact of Financial DEA due to various industry features, industries with particular features were selected. Also, the five-year period was chosen to provide a large enough sample size without being impacted by significant technology change.

Future research can be carried out by expanding the simulated tests to various accounting settings, including the diversity of accounting choices and regulations, and explain the impact of various accounting factors on Financial DEA. Also, research can expand the archival data to a longer time frame and various industries to increase the generalisability of the findings. However, issues such as heterogeneity, industry features, and technology changes still need to be considered. Also, the source of data can be expanded from secondary data to hand-collected data. Financial DEA has flourished due to the development of electronic databases and externally audited financial data from the annual reports. However, secondary data are normally highly aggregated, combining physical measures, price information, accounting information, and operational characteristics that cannot be accurately analysed. It is difficult to isolate one specific factor and test its impact using archival data, while simulated data rely on a range of interrelated assumptions that may not represent reality. Hand-collected data can provide access to specific information without losing reality. For instance, price information is commonly sensitive information for businesses. This information cannot usually be accessed from electronic databases or annual reports. The test in this study is limited to a simulation due to data accessibility. It is recommended that future studies collect specific data from the source to provide insights into how individual factors impact Financial DEA results.

7.4. Chapter Summary

The overarching goal of this study was to build a conceptual foundation for future research to apply Financial DEA. The three research questions served the overarching goal by providing a framework at the construct level (a typology) describing the application of Financial DEA, a conceptual framework at the modelling level highlighting the methodological issues, and a conceptual framework at the factor level illustrating selective magnitudes of the methodological issues. The typology structured the literature and discovered 12 dimensional constructs and nine indicators in Financial DEA. The conceptual framework at the modelling level found four quadrants of measurement models for Financial DEA and various measurement errors that could arise depending on the type of models. The conceptual framework at the factor level contains empirical tests using simulated and archival data to quantify selective measurement errors. The findings show that the Financial DEA results imply different interpretations of the findings through different reflective and formative modelling lenses. Under the view of reflective modelling, Financial DEA results diverge from technical efficiency measured by physical measures. Accounting variables include various information other than physical measures, such as prices and accounting information. However, under formative modelling, Financial DEA forms composite constructs of both physical technical efficiency and various types of information brought in by accounting variables, such as prices and accounting information, which improve the informativeness of the composite construct for benchmarking and contributing to well-informed decision making.

In conclusion, this study developed a conceptual foundation for Financial DEA by crossing several disciplines: performance measurement, management accounting, financial accounting, economics, and operational research. The conceptual foundation developed consists of three nested conceptual frameworks at the construct level, the modelling level, and the factor level, respectively. This study provides a conceptual foundation for future studies of Financial DEA in multiple research settings.

Appendix 1 Typology of Financial DEA Literature

Please find on next page

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2020	Ahn et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2010	Akdeniz et al.	benchmarking -- financial performance -- marketing efficiency						advertising expense; marketing expense; showroom occupancy expense; investment in customer					sales								
2008	Al-Sharkas et al.	productive efficiency	labour cost		PPE					core deposits	market price of purchased funds; financial equity capital						consumer loans; business loans; real estate loans;				
2017	Andreoua et al.	benchmarking -- financial performance		cost of inventory	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2018	Aparicio and Kapelko	operational efficiency	labour cost	cost of materials	PPE; gross investment in fixed assets								revenue								
2007	Asmild and Tam	profitability efficiency	remuneration					interest cost; other expenses					interest revenue; non-interest revenue								
1997	Athanassopoulos	intermediation efficiency						non-interest cost; total interest cost					non-interest income				loans		time deposit accounts; saving deposit accounts; current deposit accounts		
1995	Athanassopoulos	benchmarking -- meet the performance goal	labour cost	cost of inventory	maintenance	investments		service expenses;					governmental grants; fee and charges; rate of charges			\/	loans				
1999	Avkiran	profitability efficiency						interest expense; non-interest expense					net interest income; non-interest income								
2000	Avkiran	profitability efficiency						interest expense; non-interest expense					net interest income; non-interest income								
2009	Avkiran	profitability efficiency						interest expense; non-interest expense					net interest income; non-interest income								
2011	Avkiran	profitability efficiency						interest expense; non-interest expense					net interest income; non-interest income								
2011	Avkiran	profitability efficiency						interest expense on customer deposits; other interest expense; personnel expenses; other operating expenses					interest income on loans; other interest income; net fees and commissions; other operating income								
2018	Baghdadi et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2011	Baik et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2013	Baik et al.	benchmarking -- financial performance		COGS	NPPE			XSGA					sales								
2013	Banker et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2002	Banker et al.	operational efficiency	professional labour cost					operating costs					revenue								
2019	Banker et al.	benchmarking -- financial performance																			
2020	Banker et al.	benchmarking -- financial performance		cogs	DP			selling and administrative expenses					sales								
2019	Bao et al.	benchmarking -- financial performance		cost of inventory	PPE; operating leases		past R&D expenditures; intangibles	general and administrative expenses					sales								
2016	Benli and Bozoklu	benchmarking -- financial performance -- marketing						advertising expenditure					sales revenue								
2009	Berger et al.	intermediation efficiency				total earning assets		interest expenses to total deposits; non-interest expenses to fixed assets									total loan; liquid assets; other earning assets		total deposit		
1997	Bhattacharyya et al.	intermediation efficiency						interest expense; operating expense									loans(advances);investment		deposit		
2017	Bonsall IV et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
1999	Bowlin	benchmarking -- financial performance						operating expenses; identifiable assets					sales; operating profit; operating cash flow								

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2008	Bowlin and Renner	benchmarking -- compensation			√	assets	√	operating income; operating cash flow; CEO annual compensation; CEO long-term compensation				sales; net income								executive annual compensation; executive long-term compensa	
2008	Bowlin and Renner	benchmarking -- compensation			√	assets	√	operating income; operating cash flow; CEO annual compensation; CEO long-term compensation				sales; net income								executive annual compensation	
2008	Bowlin and Renner	benchmarking -- compensation			√	assets	√	operating income; operating cash flow; CEO annual compensation; CEO long-term compensation				sales; net income								executive long-term compensa	
1997	Brockett et al.	intermediation efficiency						interest expense; non-interest expense		provision for loan losses; total deposits			interest income; total non-interest income				total loans		allowance for loan losses		
2019	Brooks et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2013	Bryan et al.	benchmarking -- financial performance		COGS	CAPEX			selling and distribution expenses					sales revenue								
2018	Bui et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2005	Camanho and Dyson	intermediation efficiency						non-interest costs; interest costs from deposits; interest costs from loans									total loan; total value of off balance sheet business		total deposits		
2015	Chang et al.	benchmarking -- financial performance		COGS	CAPEX			selling and distribution expenses					sales revenue								
2013	Chang et al.	benchmarking -- financial performance		COGS	CAPEX			XSGA					sales								
2016	Chang et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
1990	Charnes et al.	intermediation efficiency						total operating expense; total non-interest expense; actual loan losses		provision for loan losses			total operating income; total interest income; total noninterest income				total net loans				
2011	Chen and Chen	benchmarking -- financial performance			√	total assets	√	operation cost; selling and administrative expenditure					net sales								
2019	Chen and Chen	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2019	Chen and Chu	benchmarking -- financial performance		COGS	NPPE; net operating leases		purchased goodwill; other intangibles	selling and administrative expenses; net R&D expenses					sales								
2014	Chen et al.	intermediation efficiency						operating expenses		amount of total liabilities	amount of owners' equity								incurred claims + additions to reserves	investment assets	
2015	Chen et al	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2020	Chen et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2018	Cheng et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2017	Cheung et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2020	Choo et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2017	Cornaggia et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2020	Cui and Leung	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2007	Delmas et al.	operational efficiency	labour cost	production expenses; cost of energy	PPE			transmission expenses; distribution expenses; sales, admirative and general expenses					low-voltage sales; high-voltage sales; sales foe resale								

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2018	Demerjian	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2020	Demerjian et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2012	Demerjian et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2013	Demerjian et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2014	Doumpos and Cohen	benchmarking -- meet the performance goal										fees and charges; tax revenue; subsidies from central government				book value of: recreational facilities; roads infrastructure; pavements; lighting infrastructur				cost of goods and services	
2003	Durand and Vargas	benchmarking -- financial performance			PPE		R&D expenditure	marketing expenditure; education expenditure					gross profit; sales								
2013	Dutta	operational efficiency	labour cost	cost of materials				expense related to business service			total investment		premiums earned; income from investment								
2011	Dutta and Segupta	profitability efficiency						operating expense relating to insurance business, commission paid to agents; net benefits paid to the					net premium revenue								
2008	Edirisinghe and Zhang	benchmarking -- financial performance		cost of inventory	V	account receivable; total assets	V			total liability; long-term debt			revenue; net income								
2014	Evans et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2003	Fanchon	benchmarking -- financial performance			NPPE; net investment		R&D; purchased goodwill; other intangibles	interest expense; other production expenses; advertising expense					sales								
2018	Fernandes et al.	profitability efficiency						interest expenses; operating expenses					total income								
2020	Fernando et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2001	Feroz et al.	benchmarking -- financial performance			V	assets	V	cost			common equity		revenue								imported value
2003	Feroz et al.	benchmarking -- financial performance			V	assets	V	cost			common equity		revenue								
2016	Francis et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2019	Francis et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2012	Frijns et al.	benchmarking -- financial performance		COGS	NPPE; CAPEX	TA-NPPE	V	XSGA					sales								
2012	Frijns et al.	benchmarking -- financial performance -- share market			NPPE; CAPEX			XSGA		long-term debt	book value of equity				market value of						
2009	Gaganis et al.	intermediation efficiency						interest expenses; non-interest expenses		loan loss provisions			interest income; non-interest income								
2009	Gaganis et al.	intermediation efficiency						interest expenses; non-interest expenses					interest income; non-interest income								
2020	Gao et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2020	Gao et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2014	Garrett et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2008	Giokas	productive efficiency	personnel cost					running and other operating costs					non-interest income				loan		deposits		
2008	Giokas	productive efficiency	personnel cost					running and other operating costs									loan transactions		deposit transitions		remain transition
2008	Giokas	profitability efficiency						interest costs; non-interest costs					interest income; non-interest income								
2015	Giokas et al.	benchmarking -- financial performance			V	total assets	V	operating cost					total sales								

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2018	Gul	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2016	Guo	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2009	Guzman et al.	operational efficiency	staff costs	production expenses	DP			other operating expenses					revenue								
2008	Guzman and Arcas	operational efficiency	staff costs	cost of materials	DP			other operating expenses					revenue								
2008	Guzman and Arcas	operational efficiency	staff costs		PPE								revenue								
2011	Hadad et al.	intermediation efficiency	employee expenses					non-employee expenses		consumer deposits and commercial borrowing; provisions			off-balance-sheet income				commercial loans; other earning assets				
2019	Hapter et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2018	Harper and Sun	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2020	Hasan	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2009	Ho et al.	benchmarking -- financial performance			\/	total assets	\/	operating expense			total equity		revenue; gross profit								
2009	Ho et al.	benchmarking -- financial performance -- share market										revenue; gross profit	net income		EPS						
2014	Hsiao	benchmarking -- financial performance						operating costs; operating expense; non-business expenditure					non-business income; net business income								
2013	Hsiao and Lin	intermediation efficiency						interest expense; non-interest expense		total deposit			interest revenue; non-interest revenue				total loans				
2010	Hsiao et al.	intermediation efficiency						interest expense; non-interest expense		total deposit			interest revenue; non-interest revenue				total loans				
2019	Hsu and Khan	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2008	Huang et al.	intermediation efficiency						interest expense; non-interest expense		total deposit			interest revenue; non-interest revenue				total loans				
2007	Joo et al.	operational efficiency	payroll		occupancy expenses			other expenses					revenue-textiles; revenue-wares; other revenues								
2011	Joo et al.	benchmarking -- financial performance -- total assets			PPE	current assets; other assets	\/						revenue								
2011	Joo et al.	benchmarking -- financial performance -- current assets				cash and cash equivalent; account receivable;							revenue								
2011	Joo et al.	operational efficiency		COGS	DP			XSGA					revenue								
2009	Joo et al.	operational efficiency	wages and benefits	COGS				other expenses; occupancy expenses					sales								
2009	Joo et al.	operational efficiency	wages and benefits	COGS				other expenses					sales- restaurant; sales-retail								
2010	Joo et al.	operational efficiency		COGS	DP			XSGA					revenue								
2010	Joshi and Singh	operational efficiency	wages and salaries	cost of materials; cost of energy	NPPE								gross sales								
2013	Jung et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2018	Kaffash et al.	intermediation efficiency			PPE					deposits & short term funding	equity		net income				loans				
2018	Kaffash et al.	profitability efficiency						interest expense; non-interest expense					interest income; non-interest income								
2008	Kao and Hwang	profitability efficiency						operating expense; insurance expense					direct written premiums; reinsurance premiums								
2008	Kao and Hwang	profitability efficiency										direct written premiums; reinsurance premiums	underwriting profit; investment profit								
2004	Kao and Liu	intermediation efficiency						interest expense; non-interest expense		total deposit			interest income; non-interest income				loans				
2017	Kapelko	operational efficiency	labour cost	cost of materials	PPE								revenue								
2017	Kapelko	operational efficiency	labour cost	cost of materials	PPE; gross investment in fixed assets								revenue								

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2018	Kapelko	operational efficiency	labour cost	cost of materials	PPE								operating revenue								
2017	Kapelko and Lansink	operational efficiency	labour cost	cost of materials	DP;PPE; GPPE								revenue								
2018	Kapelko and Lansink	operational efficiency	labour cost	cost of materials	DP;PPE; GPPE								revenue								
2015	Kapelko and Lansink	operational efficiency	labour cost	cost of materials	PPE								operating revenue								
2014	Kapelko et al.	operational efficiency	labour cost	cost of materials	PPE								sales + change in the stock								
2015	Kapelko et al.	operational efficiency	labour cost	cost of materials	PPE								sales + change in the stock								
2017	Kapelko et al.	operational efficiency	labour cost	cost of materials	DP; gross investment in fixed assets								revenue								
2016	Kapelko et al.	operational efficiency	labour cost	cost of materials	PPE; gross investment in fixed assets								sales + change in the stock								
2017	Kapelko et al.	operational efficiency	labour cost	cost of materials	PPE								sales + change in the stock								
2015	Kapelko et al.	operational efficiency	labour cost	cost of materials	PPE								sales + change in the stock								
2015	Kapelko et al.	operational efficiency	labour cost	cost of materials	PPE								operating revenue								
2017	Koester et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2008	Kown et al.	benchmarking -- financial performance -- total assets			PPE	current assets; other assets	√						revenue								
2008	Kown et al.	benchmarking -- financial performance -- current assets				cash and cash equivalent; account receivable; inventory; other current assets							revenue								
1998	Kozmetsky and Yue	benchmarking -- financial performance		COGS	√	total assets	√	XSGA					net sales								
2015	Krishnan and Wang	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2020	Kuang et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2016	Kubick and Lockhart	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2017	Kuo et al.	operational efficiency	staff costs					operating expense							incurred claims in addition to reserves		investment assets				
2017	Kuo et al.	intermediation efficiency				investment assets		incurred claims			addition to reserves		net income; operating cashflow								
2019	Kuo et al.	benchmarking -- financial performance			PPE		R&D expense	selling expense					EBIT								
2019	Kuo et al.	benchmarking -- financial performance							increase in long-term debt	total long-term debts	capital surplus; total equity	EBIT	net income								interest payment; tax payment
2020	Kuo et al.	benchmarking -- financial performance -- share market										free cash flow			market value						cash dividend
2014	Kweh et al.	benchmarking -- financial performance						operating expense	√	liabilities	shareholders' equity		revenue		market value						
2018	Kweh et al.	productive efficiency	staff costs		PPE			other expenses									other earning assets				
2018	Kweh et al.	intermediation efficiency				other earning assets; loans; investment assets				deposits			profit; interest income								
2015	Kweh et al.	benchmarking -- financial performance	staff costs	COGS				operating expense					sales		market value						
2014	Kweh et al.	intermediation efficiency						operating expense		total liability	owners' equity		investments profits		incurred loss plus addition to						
2021	Kweh et al.	profitability efficiency	staff expense					administrative expenses; selling expense					net sales		market value						
2008	Kwon et al.	benchmarking -- financial performance			DP			cost of service; XSGA; depreciation and amortisation					revenue								
2005	Lee	benchmarking -- financial performance						interest expense; operating expense					sales								
1998	Leightner and Lovell	productive efficiency	personnel cost		PPE					provision for loan loss			interest income; non-interest income								

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2018	Liu et al.	benchmarking -- financial performance						room expenses; food and beverage expense; administrative and general expense; marketing expense; other expense					room revenues; food and beverage revenue; other revenues								
2014	Lu et al.	intermediation efficiency						operating expense			total capital		investments profits		incurred benefits plus addition to						
2005	Luo and Donthu	benchmarking -- financial performance -- marketing						print expenses; broadcast expenses; outdoor					sales								
2009	Magdaleno Garcia-Garcia	operational efficiency	staff costs	cost of materials	NPPE			operating expenses					operating revenue								
2015	Mahajan et al.	operational efficiency	wages and salaries	cost of materials	GPPE			adverting and marketing expense					net sales								
2018	Mahajan et al.	operational efficiency	wages and salaries	cost of materials	GPPE			adverting and marketing expense					net sales								
2014	Mahajan et al.	operational efficiency	wages and salaries	cost of materials	GPPE			adverting and marketing expense					net sales								
2014	Mahajan et al.	operational efficiency	wages and salaries	cost of materials	GPPE			adverting and marketing expense					net sales								
1999	Majumdar	operational efficiency	employee expenses	cost of materials	capital usage cost			expenditure on other operating items					gross output								
2002	Manandhar and Tang	profitability efficiency						interest expense; non-interest expense					interest revenue; non-interest revenue								
2013	Mazumdar	operational efficiency	wages and salaries	cost of materials; cost of energy	NPPE								sales								
2009	Mazumdar and Rajeev	operational efficiency	wages and salaries	cost of materials; cost of energy	PPE								sales + change in the stock								
2007	McEachern and Paradi	profitability efficiency	remuneration					interest expense; other expenses					interest revenue; non-interest revenue								
2018	Mckee Kagan	intermediation efficiency	commissions		PPE					deposits							loans and investments				off balance sheet activities
2012	Mhatre et al.	operational efficiency		cost of inventory ; COGS				XSGA					revenue								
1996	Millter and Noulas	intermediation efficiency						interest expense; non-interest expenses		total transactions deposits; total non-transactions deposits			interest income; non-interest income				commercial and industrial loans; consumer loans; real estate loans; investments				
2006	Min and Joo	operational efficiency	wages and salaries		PPE	account receivables		operating expenses					operating income								
2009	Min and Joo	benchmarking -- financial performance		COGS	DP			SGA;depreciation and amortisation					revenue								
2009	Min and Joo	benchmarking -- financial performance -- total assets			PPE	current assets; other assets	\/						revenue								
2007	Min et al.	operational efficiency	payroll and labour related expenses	COGS				other operating and non-operating expenses					revenue from rooms; revenue from food and beverage; revenue from other services								
2007	Min et al.	benchmarking -- financial performance						expenses for rooms; expenses for food and beverage; expenses associated with other services					revenue from rooms; revenue from food and beverage; revenue from other services								
2008	Noulas et al.	productive efficiency	personnel cost					other operating expenses									loans; other loans		deposits		financial products
2018	Nourani et al.	profitability efficiency	staff costs					management expenses					premium earned; operating cashflow								
2018	Nourani et al.	benchmarking -- financial performance			PPE										net claims plus additions to reserve		total investments				
2018	Nourani et al.	benchmarking -- financial performance--funding efficiency									equities		investment income; net profit								
2017	Nourani et al.	operational efficiency	management expenses; staff cost			total assets							total premiums earned								
2019	Nourani et al.	intermediation efficiency	staff costs			total earning assets; other assets		total administrative and management expenses					net profit; net interest income								
2010	Oberholzer	intermediation efficiency			PPE						equity						loans		deposits		

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2012	Oberholzer	benchmarking -- financial performance			PPE			total expenditure			book value of shareholder equity		sales		market value of equity					dividend pay-out	
2013	Oberholzer	benchmarking -- financial performance		COGS				other expenses					revenue; EBITDA								
2014	Oberholzer	benchmarking -- financial performance--funding efficiency									equity					V	assets				
2014	Oberholzer	benchmarking -- financial performance -- total assets			V	assets	V						revenue								
2014	Oberholzer	benchmarking -- financial performance										revenue	profit								
2014	Oberholzer	benchmarking -- compensation	remuneration										revenue		total equity	market value of assets				total cost	
2009	Oberholzer and Van der Westhuizen	intermediation efficiency			PPE	investments					financial capital						loans		deposits		
2010	Oberholzer and Van der Westhuizen	productive efficiency	staff costs					operating cost		deposits			interest income; non-interest income								
2010	Oberholzer and Van der Westhuizen	productive efficiency			PPE	financial capital										V	loans		deposits		
2012	Oberholzer and Theunissen	benchmarking -- compensation	compensation										profit			V	assets				
2017	Oberholzer et al.	benchmarking -- financial performance--funding efficiency							current liability (t-1)		capital (t-1)		sales; net profit; NOPAT								
2017	Ochola	intermediation efficiency				total assets		commission and other expenses			shareholders capital life fund and reserve		net earned premium; investment income and other incomes; profit and loss after tax							net incurred claims	
1990	Oral and Yolalan	profitability efficiency	personnel cost		DP			administrative expense; interest paid on deposits					interest income; non-interest income								
1992	Oral et al.	profitability efficiency	personnel cost		DP			administrative expense; interest paid on deposits; non-interest expense					interest income; non-interest income								
2011	Paradi et al.	intermediation efficiency			PPE	cash; non-performing loans; wealth management		loan loss		other liabilities							homeowner mortgage; commercial loans		consumer lending; commercial deposits; consumer deposits		
2011	Paradi et al.	profitability efficiency	employee expenses		occupancy expenses			loan loss; cross charges; other expenses; sundry					commissions				wealth management; home mortgages; commercial loans		consumer deposits; consumer landing; commercial deposits;		
1999	Parkan and Wu	operational efficiency	compensation	cost of materials	CAPEX			other expenses					sales of goods, industrial work and industrial services; other receipts; disposal of fixed assets								
2020	Phan et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2000	Piesse and Thirtle	operational efficiency	wages	cost of materials; cost of energy	PPE								gross output value								
2002	Pille and Paradi	intermediation efficiency						non-interest expense		deposits			interest income and other incomes		equity		loans, cash and investments				
2002	Pille and Paradi	intermediation efficiency						non-interest expense		deposits			interest income and other incomes		equity		loans; cash and investments				
2002	Pille and Paradi	intermediation efficiency						interest expense; non-interest expense					interest income and other incomes				loans, cash and investments		deposits		
2014	Qiu et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2008	Raheman et al.	benchmarking -- financial performance		COGS	V	total assets	V	operating expense			shareholders' equity		sales								
2019	Rahman et al.	benchmarking -- financial performance -- marketing efficiency						TV advertising; radio advertising; magazine advertising; newspaper advertising; outdoor advertising; online advertising					sales								

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2016	Rahman et al.	benchmarking -- financial performance -- marketing efficiency			expenditure on bank branch premises			advertising expenditure								√	total assets		total deposit		
2018	Rahman et al.	benchmarking -- financial performance -- marketing				account receivable		advertising expenditure: R&D expenditure										brand value			
2011	Rodriguez-Peres et al.	operational efficiency			land and building (historical based and fair-value based)	financial investments in associated and group companies(historical based and fair-value based); other financial investments (historical based		total expenses					revenue								
2017	Sallehu	benchmarking -- financial performance		COGS	CAPEX			XSGA; CAPX					sales								
2009	Saranga	operational efficiency	labour cost	cost of materials	CAPEX			sundry expense					gross income								
2009	Saranga and Phani	operational efficiency	wages and salaries	cost of materials				cost of production and selling					net sales								
2019	Sarwar et al.	benchmarking -- financial performance		COGS	PPE		goodwill, other intangibles	selling expenses; finance expenses; administrative expenses; intangibles; R&D expenses					sales								
2003	Sathye	profitability efficiency						interest expense; non-interest expense					interest income; non-interest income								
2000	Sathye	intermediation efficiency	labour cost		CAPEX	loanable funds											loans		deposits		
2017	Sathye and Sathye	profitability efficiency						interest expense; non-interest expense					interest income; non-interest income								
2001	Severovic et al.	productive efficiency	salaries			credits granted		banking expenditure; operating expenditure					credit profits; banking profits						deposits		
2020	Shin and Park	benchmarking -- financial performance		COGS	PPE		intangible assets	XSGA					sales								
2002	Shiu	operational efficiency	remuneration	intermediate inputs	NPPE								gross industrial output								
1990	Smith	benchmarking -- financial performance--funding efficiency						interest expense; tax payment	√	average debts	average equity			net income							
2012	Sohob et al.	operational efficiency	remuneration	cost of materials	GPPE								sales								
2017	Sun	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2018	Tone et al.	operational efficiency			PPE			operating expenses							incurred claims + additions to reserves						
2018	Tone et al.	intermediation efficiency				(with and without) investment assets					incurred claims + additions to reserves		premiums earned; (with and without) income from investment								
2020	Truong et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2011	Tsolas	benchmarking -- financial performance						operating expenses, selling and administrative cost					revenue								
2011	Tsolas	benchmarking -- financial performance										revenue	net income before tax								
2015	Walters et al.	benchmarking -- financial performance		COGS	NPPE; capitalised operating lease;		capitalised R&D investments; acquired good will; intangible assets	XSGA					revenues								
2006	Wang	operational efficiency		cost of production				operating expenses					sales								
2020	Wang	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2014	Wang et al.	benchmarking -- financial performance						operating expenses	√	liabilities	stockholder equity		revenues		market value			intangible assets			
2016	Wang et al.	benchmarking -- financial performance	staff costs		NPPE		R&D costs						revenues				software assets				
2016	Wang et al.	benchmarking -- financial performance -- share market				software assets						revenues	total profits		market value						
2017	Wang et al.	benchmarking -- financial performance						operating expenses	√	liabilities	stockholder equity		revenue		market value			intangible assets			
2018	Wang et al.	benchmarking -- financial performance						operating expense	√	total liabilities	stockholder equity		revenue		market value			intangible assets			

Year	Authors	Firm Performance Construct	Inputs										Outputs								
			Labour	Material	Capital	Current Assets	Intangible Assets	Expenses	Current Liability	Non-current Liability	Equity	Revenue	Revenue	Net Income	Equity	Non-current Assets	Assets	Intangible	Liability	Expenses	Other
2017	Wang et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
2017	Wang et al.	benchmarking -- financial performance					intangible assets	operating expenses	V	liabilities	equities		revenue		market value						
2020	Wang et al.	benchmarking -- financial performance		COGS	NPPE		R&D expenses, goodwill, other intangibles	XSGA					operating profit	net income							
2020	Wanke et al.	benchmarking -- financial performance	personnel cost			net loan							net interest margin								
2020	Wanke et al.	benchmarking -- financial performance			PPE	total earning assets									total equity						
2020	Wanke et al.	benchmarking -- financial performance						costs			loan loss reserve		net income								
2020	Wen et al.	benchmarking -- financial performance		COGS	NPPE; net operating leases		R&D, goodwill, other intangibles	XSGA					sales								
1998	Worthington	benchmarking -- financial performance			PPE	current assets		operating expenses					sales; profit before tax		market capitalisatio						
2004	Worthington	operational efficiency			PPE			interest and non-interest expenses		at-call deposits; notice-of-withdrawal; deposits; fixed term deposits			interest and non-interest income				personal loan; commercial loan; residential loans; investments				
2002	Worthington and Hurley	operational efficiency	remuneration		CAPEX			information technology expense; interest expenses					premium				invested assets				
2016	Wu et al.	operational efficiency			NPPE			operating expenses		liabilities	stockholder equity				incurred claims + additions to reserves		investment assets				
2016	Wu et al.	intermediation efficiency				investment assets					incurred claims + additions to reserves		operating income; operating cash flow								
2016	Wu et al.	productive efficiency	personnel cost		PPE	liquid assets		other operating expenses									loans; other earning assets		deposits		
2016	Wu et al.	intermediation efficiency				loans; other earning assets				deposits			net interest income								
2014	Yang and Choi	benchmarking -- financial performance				cash and cash equivalents; inventory; PPE							net sales								
2014	Yang and Choi	benchmarking -- financial performance		COGS				operating expenses					net sales								
2013	Yang et al.	benchmarking -- financial performance	remuneration	COGS	PPE			operating expenses					sales		market value						
1996	Yeh	intermediation efficiency						interest expense; non-interest expenses		deposits			interest income; non-interest income				loans				
2014	Yu et al.	benchmarking -- financial performance				intangible resources		XSGA; relationship expenditure					sales								
2014	Yu et al.	operational efficiency	labour cost		GPPE								cost of sales								
2010	Yusof et al.	benchmarking -- financial performance			V	total assets	V	operating expenses					sales								
2006	Zelenyuk and Zheka	operational efficiency	remuneration		DP			operating expenses					sales								
2006	Zhang et al.	benchmarking -- financial performance	compensation					non-labour non-interest expense			equity		commission revenue; trading gains from market making; investment banking revenue; revenue from asset management; total								
2008	Zhou et al.	operational efficiency	remuneration		NPPE			operating expenses		current liability			sales								

Appendix 2 Test One Descriptive Statistics

(This table is an extension of Table 5 - 2)

Panel A: N = 24, iteration = 100									
		MEAN	**SD**	**CV**	**MIN**	**Q1**	**MED**	**Q3**	**Max**
Physical	X_1	7.50	0.74	0.10	4.96	7.01	7.52	8.04	10.19
	X_2	7.49	0.73	0.10	5.15	6.99	7.50	7.99	10.04
	Y	7.48	0.53	0.07	5.68	7.11	7.47	7.85	9.08
	Y_{adj}	6.41	0.88	0.14	3.57	5.85	6.46	7.03	8.91
Constant	W_{x1c}	10.00	0.00	0.00	10.00	10.00	10.00	10.00	10.00
	W_{x2c}	15.00	0.00	0.00	15.00	15.00	15.00	15.00	15.00
	W_{yc}	30.00	0.00	0.00	30.00	30.00	30.00	30.00	30.00
	X_1W_{x1c}	75.05	7.40	0.10	49.58	70.10	75.20	80.41	101.92
	X_2W_{x2c}	112.37	10.95	0.10	77.24	104.82	112.50	119.80	150.63
	$Y_{adj}W_{yc}$	192.33	26.39	0.14	106.97	175.62	193.79	210.85	267.16
Narrow	W_{x1n}	9.99	0.50	0.05	8.35	9.65	10.00	10.34	11.68
	W_{x2n}	14.98	0.76	0.05	12.57	14.48	14.99	15.50	18.23
	W_{yc}	30.04	1.52	0.05	24.98	29.01	30.03	31.05	34.92
	X_1W_{x1n}	74.98	8.23	0.11	49.87	69.48	74.94	80.58	108.40
	X_2W_{x2n}	112.22	12.32	0.11	74.15	103.72	112.04	120.20	153.85
	$Y_{adj}W_{yn}$	198.17	28.54	0.14	99.95	179.99	199.43	218.23	289.46
Broad	W_{x1b}	9.98	2.45	0.25	3.06	8.37	10.01	11.63	17.73
	W_{x2b}	14.99	3.72	0.25	2.14	12.48	14.95	17.50	26.40
	W_{yb}	30.03	7.49	0.25	5.10	25.12	29.96	34.95	56.90
	X_1W_{x1b}	74.92	19.89	0.27	21.54	61.62	74.42	88.14	148.35
	X_2W_{x2b}	112.29	29.94	0.27	14.26	91.85	112.00	132.19	211.31
	$Y_{adj}W_{yb}$	198.12	56.61	0.29	36.07	159.07	195.52	235.04	411.22
Panel B: N = 96, iteration = 100									
		MEAN	**SD**	**CV**	**MIN**	**Q1**	**MED**	**Q3**	**Max**
Physical	X_1	7.51	0.74	0.10	4.38	7.01	7.50	8.01	10.16
	X_2	7.50	0.75	0.10	4.77	6.99	7.49	8.01	10.57
	Y	7.48	0.53	0.07	5.51	7.12	7.49	7.84	9.58
	Y_{adj}	6.42	0.86	0.13	3.25	5.84	6.47	7.04	9.21
Constant	W_{x1c}	10.00	0.00	0.00	10.00	10.00	10.00	10.00	10.00
	W_{x2c}	15.00	0.00	0.00	15.00	15.00	15.00	15.00	15.00
	W_{yc}	30.00	0.00	0.00	30.00	30.00	30.00	30.00	30.00
	X_1W_{x1c}	75.06	7.40	0.10	43.79	70.06	75.03	80.06	101.64
	X_2W_{x2c}	112.44	11.29	0.10	71.56	104.85	112.30	120.09	158.50
	$Y_{adj}W_{yc}$	192.72	25.82	0.13	97.54	175.29	194.23	211.13	276.16
Narrow	W_{x1n}	10.00	0.50	0.05	8.29	9.66	10.01	10.33	12.00
	W_{x2n}	14.99	0.76	0.05	12.03	14.49	14.99	15.49	18.11
	W_{yc}	30.00	1.51	0.05	24.42	28.97	30.00	31.01	35.63

	X_1W_{x1n}	75.03	8.26	0.11	41.80	69.29	74.98	80.60	108.24
	X_2W_{x2n}	112.37	12.59	0.11	63.46	103.91	112.08	120.71	164.81
	$Y_{adj}W_{yn}$	195.06	27.99	0.14	97.80	176.01	196.15	214.57	281.42
Broad	W_{x1b}	10.00	2.48	0.25	0.51	8.35	10.00	11.69	19.53
	W_{x2b}	14.97	3.73	0.25	1.74	12.44	14.94	17.48	28.85
	W_{yb}	30.09	7.45	0.25	2.48	25.01	30.01	35.14	59.55
	X_1W_{x1b}	75.10	20.31	0.27	3.91	61.25	74.67	88.29	167.02
	X_2W_{x2b}	112.16	30.08	0.27	12.27	91.47	111.07	131.69	240.56
	$Y_{adj}W_{yb}$	195.62	55.29	0.28	9.19	157.41	192.74	230.99	457.83

Note: (a) X_1, the first physical input. X_2, the second physical input. Y, the physical output, calculated as $Y = X_1^{0.5}X_2^{0.5}$. Y_{adj}, the adjusted physical output, with inefficiencies. $W_{x_{1c}}$, the price of the first input in the constant price scenario. $W_{x_{2c}}$, the price of the second input in the constant price scenario. W_{yc}, the price of the output in the constant price scenario. $W_{x_{1n}}$, the price of the first input in the narrow price scenario. W_{x_2n}, the price of the second input in the narrow price scenario. W_{yn}, the price of the output in the narrow price scenario. $W_{x_{1b}}$, the price of the first input in the broad price scenario. $W_{x_{2b}}$, the price of the second input in the broad price scenario. W_{yb}, the price of the output in the broad price scenario. (b) The descriptive features include minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), the maximum value (MAX), the mean value (MEAN), standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (c) This table extends Table 5 - 2.

Appendix 3 Test One Efficiency Scores

(This table extends Table 5 - 3)

Panel A: N = 24, iteration = 100										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Physical	0.8821	3.36	4.16	0.5071	0.4929	0.8157	0.8990	0.9730	0.1019	0.1155
Constant	0.8821	3.36	4.16	0.5071	0.4929	0.8157	0.8990	0.9730	0.1019	0.1155
Narrow	0.8459	2.68	3.04	0.5717	0.4283	0.7683	0.8587	0.9365	0.1130	0.1336
Broad	0.6438	2.59	2.74	0.9123	0.0877	0.4809	0.6283	0.8015	0.2148	0.3336
Panel B: N = 96, iteration = 100										
	MEAN	**100%**	**99%**	**RANGE**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Physical	0.8690	4.85	8.73	0.5246	0.4754	0.8042	0.8843	0.9502	0.0983	0.1132
Constant	0.8690	4.85	8.73	0.5246	0.4754	0.8042	0.8843	0.9502	0.0983	0.1132
Narrow	0.8070	3.19	4.00	0.5981	0.4019	0.7336	0.8145	0.8858	0.1079	0.1338
Broad	0.5278	2.94	3.03	0.9784	0.0216	0.3879	0.5030	0.6392	0.1911	0.3621

Note: (a) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs that are at least 99% efficient (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (b) All efficiency scores are generated from the CRS model. (c) This table extends Table 5 - 3.

Appendix 4 Test Two Variable Definitions

Variables	Definition
AGE_B	Age of property plant and equipment, broad distribution
AGE_C	Age of property plant and equipment when the value is constant
AGE_N	Age of property plant and equipment, narrow distribution
AGE_RATE_B	Rate of age to useful life, broad distribution
AGE_RATE_C	Rate of age to useful life, constant
AGE_RATE_N	Rate of age to useful life, narrow distribution
BOOT_BOX_AGE	Age of property plant and equipment for the box industry based on the bootstrapped data. Assuming no residual value, the straight-line method
BOOT_BOX_DP	Bootstrapped depreciation expense value of the box industry
BOOT_BOX_GPPE	The bootstrapped gross value of property plant and equipment for the box industry
BOOT_BOX_NPPE	The bootstrapped net value of property plant and equipment for the box industry
BOOT_BOX_OPEX	Bootstrapped operating expense value of the box industry
BOOT_BOX_R	Relative age rate of property plant and equipment for the box industry in the bootstrapped data. Assuming no residual value, the straight-line method
BOOT_BOX_SALES	Bootstrapped sale value of the box industry
BOOT_BOX_UL	The useful life of property plant and equipment for the box industry in the bootstrapped data. Assuming no residual value, the straight-line method
BOOT_GOLD_AGE	Age of property plant and equipment for the gold industry in the bootstrapped data. Assuming no residual value, the straight-line method
BOOT_GOLD_DP	Bootstrapped depreciation expense value of the gold industry
BOOT_GOLD_GPPE	The bootstrapped gross value of property plant and equipment for the gold industry
BOOT_GOLD_NPPE	The bootstrapped net value of property plant and equipment for the gold industry
BOOT_GOLD_OPEX	Bootstrapped operating expense value of the gold industry
BOOT_GOLD_R	Relative age of property plant and equipment for the gold industry in the bootstrapped data. Assuming no residual value, the straight-line method
BOOT_GOLD_SALES	Bootstrapped sale value of the gold industry

BOOT_GOLD_UL	The useful life of property plant and equipment for the gold industry in bootstrapped data. Assuming no residual value, the straight-line method
BOX_AGE	Age of property plant and equipment for the box industry. Assuming no residual value, the straight-line depreciation method
BOX_DP	Depreciation expense value of the box industry inflated to the currency value of FY2019
BOX_GPPE	The gross value of property plant and equipment for the box industry inflated to the currency value of FY2019
BOX_NPPE	The net value of property plant and equipment for the box industry inflated to the currency value of FY2019
BOX_OPEX	Operating expense value of the box industry, inflated to the currency value of FY2019
BOX_R	Relative age of property plant and equipment for the box industry. Assuming no residual value, the straight-line depreciation method
BOX_SALES	The sale value of the box industry inflated to the currency value of FY2019
BOX_UL	The useful life of property plant and equipment for the box industry. Assuming no residual value, the straight-line depreciation method
DP_AGE_B	Depreciation expense value when the factor of age varies around a broad distribution
DP_AGE_N	Depreciation expense value when the factor of age varies around a narrow distribution
DP_BOTH_B	Depreciation expense value, when both the factor of useful life and the factor of age vary around a broad distribution
DP_BOTH_N	Depreciation expense value, when both the factor of useful life and the factor of age vary around a narrow distribution
DP_UL_B	Depreciation expense value when the factor of useful life varies around a broad distribution
DP_UL_N	Depreciation expense value when the factor of useful life varies around a narrow distribution
GOLD_AGE	Age of property plant and equipment for the gold industry. Assuming no residual value, the straight-line depreciation method
GOLD_DP	Depreciation expense value of the gold industry inflated to the currency value of FY2019
GOLD_GPPE	The gross value of property plant and equipment for the gold industry inflated to the currency value of FY2019
GOLD_NPPE	The net value of property plant and equipment for the gold industry inflated to the currency value of FY2019
GOLD_OPEX	Operating expense value of the gold industry, inflated to the currency value of FY2019
GOLD_R	Relative age of property plant and equipment for the gold industry. Assuming no residual value, the straight-line depreciation method

GOLD_SALES	The sale value of the gold industry inflated to the currency value of FY2019
GOLD_UL	The useful life of property plant and equipment for the gold industry. Assuming no residual value, the straight-line depreciation method
INEFFICIENCY	Simulated value of inefficiency
NPPE_AGE_B	The net value of property plant and equipment, when the factor of age varies around a broad distribution
NPPE_AGE_N	The net value of property plant and equipment, when the factor of age varies around a narrow distribution
NPPE_BOTH_B	The net value of property plant and equipment, when both the factor of useful life and the factor of age vary around a broad distribution
NPPE_BOTH_N	The net value of property plant and equipment, when both the factor of useful life and the factor of age vary around a narrow distribution
NPPE_UL_B	The net value of property plant and equipment, when the factor of useful life varies around a broad distribution
NPPE_UL_N	The net value of property plant and equipment, when the factor of useful life varies around a narrow distribution
SCORE_DP_AGE_B	Efficiency scores degenerated by simulation when DP proxies the capital, and the factor of AGE varies around a broad distribution
SCORE_DP_AGE_N	Efficiency scores in simulation, when DP proxies the capital, and the factor of AGE varies around a narrow distribution
SCORE_DP_BOTH_B	Efficiency scores in simulation, when DP proxies the capital, and BOTH the factor of UL and AGE vary around a broad distribution
SCORE_DP_BOTH_N	Efficiency scores in simulation, when DP proxies the capital, and BOTH the factor of UL and AGE vary around a narrow distribution
SCORE_DP_UL_B	Efficiency scores in simulation, when DP proxies the capital, and the factor of UL varies around a broad distribution
SCORE_DP_UL_N	Efficiency scores in simulation, when DP proxies the capital, and the factor of UL varies around a narrow distribution
SCORE_NPPE_AGE_B	Efficiency scores in simulation, when NPPE proxies the capital, and the factor of AGE varies around a broad distribution
SCORE_NPPE_AGE_N	Efficiency scores in simulation, when NPPE proxies the capital, and the factor of AGE varies around a narrow distribution
SCORE_NPPE_BOTH_N	Efficiency scores in simulation, when NPPE proxies the capital, and BOTH the factor of UL and AGE vary around a narrow distribution

SCORE_NPPE_BOTH_N	Efficiency scores in simulation, when NPPE proxies the capital, and BOTH the factor of UL and AGE vary around a broad distribution
SCORE_NPPE_UL_B	Efficiency scores in simulation, when NPPE proxies the capital, and the factor of UL varies around a broad distribution
SCORE_NPPE_UL_N	Efficiency scores in simulation, when NPPE proxies the capital, and the factor of UL varies around a narrow distribution
SCORES_BOX_DP	Efficiency scores in the empirical test, for box industry, DP proxies the capital
SCORES_BOX_GPPE	Efficiency scores in the empirical test, for box industry, NPPE proxies the capital
SCORES_BOX_NPPE	Efficiency scores in the empirical test, for box industry, NPPE proxies the capital
SCORES_GOLD_DP	Efficiency scores in the empirical test, for the gold industry, DP proxies the capital
SCORES_GOLD_GPPE	Efficiency scores in the empirical test, for the gold industry, GPPE proxies the capital
SCORES_GOLD_NPPE	Efficiency scores in the empirical test, for the gold industry, NPPE proxies the capital
SIMU_ADJ_SALES	Simulated value of sales adjusted by the inefficiency
SIMU_GPPE	Simulated value of property plant and equipment
SIMU_OPEX	Simulated value of operating expense
SIMU_SALES	Simulated value of sales via Cobb-Douglas function
UL_B	The useful life of property plant and equipment, broad distribution
UL_C	The useful life of property plant and equipment when the value is constant
UL_N	The useful life of property plant and equipment, narrow distribution

Appendix 5 Test Two Descriptive Statistics of Simulated Data

(This table is an extension of Table 5 - 7)

Panel A: N = 24, iteration = 100									
	Variables	**MIN**	**Q1**	**Median**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**
	SIMU_GPPE	66.29	92.93	99.78	106.49	131.19	99.66	9.80	0.10
	SIMU_OPEX	65.56	93.27	100.04	107.03	139.16	100.00	10.13	0.10
Base	*SIMU_SALES*	74.73	94.70	99.66	104.34	121.04	99.58	7.05	0.07
	INEFFICIENTCY	0.00	0.06	0.14	0.23	0.69	0.16	0.12	0.75
	SIMU_ADJ_SALES	46.36	77.42	86.07	93.97	118.84	85.39	11.42	0.13
	UL_C	10.00	10.00	10.00	10.00	10.00	10.00	0.00	0.00
	UL_N	10.33	12.07	12.49	12.94	14.71	12.53	0.64	0.05
	UL_B	10.00	10.60	12.44	14.96	48.09	13.47	3.98	0.30
	AGE_C	8.00	8.00	8.00	8.00	8.00	8.00	0.00	0.00
Factors	*AGE_N*	6.80	7.73	8.00	8.28	9.69	8.00	0.41	0.05
	AGE_B	2.08	6.69	8.04	9.43	10.00	7.89	1.71	0.22
	AGE_RATE_C	0.80	0.80	0.80	0.80	0.80	0.80	0.00	0.00
	AGE_RATE_N	0.68	0.77	0.80	0.83	0.97	0.80	0.04	0.05
	AGE_RATE_B	0.21	0.67	0.80	0.94	1.00	0.79	0.17	0.22
P2a	*DP_UL_C*	6.63	9.29	9.98	10.65	13.12	9.97	0.98	0.10
DP	*DP_UL_N*	5.29	7.39	7.97	8.57	10.82	7.98	0.87	0.11
impacted by UL	*DP_UL_B*	1.89	6.55	7.96	9.28	12.58	7.87	1.89	0.24
P2b	*NPPE_UL_C*	13.26	18.59	19.96	21.30	26.24	19.93	1.96	0.10
NPPE	*NPPE_UL_N*	2.46	16.72	19.76	22.90	37.07	19.91	4.58	0.23
impacted by UL or AGE	*NPPE_UL_B*	0.00	5.50	19.52	32.67	85.87	21.01	17.15	0.82
P2c	*NPPE_BOTH_C*	13.26	18.59	19.96	21.30	26.24	19.93	1.96	0.10
NPPE	*NPPE_BOTH_N*	4.85	30.46	35.50	40.76	62.36	35.68	7.51	0.21
impacted by UL & AGE	*NPPE_BOTH_B*	0.00	10.71	35.07	54.37	106.83	34.69	25.73	0.74

Panel B: N = 96, iteration = 100									
	Variables	**MIN**	**Q1**	**Median**	**Q3**	**MAX**	**MEAN**	**SD**	**CV**
	SIMU_GPPE	65.39	93.01	99.89	106.73	136.73	99.85	10.09	0.10
	SIMU_OPEX	61.09	93.42	100.16	106.81	142.67	100.10	9.98	0.10
Base	*SIMU_SALES*	74.44	94.95	99.69	104.48	125.04	99.72	7.10	0.07
	INEFFICIENTCY	0.00	0.06	0.13	0.23	0.76	0.16	0.12	0.76
	SIMU_ADJ_SALES	42.65	78.15	86.37	94.03	120.49	85.68	11.50	0.13
	UL_C	10.00	10.00	10.00	10.00	10.00	10.00	0.00	0.00
	UL_N	10.45	12.08	12.50	12.93	15.23	12.53	0.64	0.05
	UL_B	10.00	10.69	12.51	14.99	48.62	13.56	3.98	0.29
	AGE_C	8.00	8.00	8.00	8.00	8.00	8.00	0.00	0.00
Factors	*AGE_N*	6.57	7.73	8.00	8.28	9.57	8.00	0.40	0.05
	AGE_B	2.06	6.67	8.00	9.35	10.00	7.84	1.72	0.22
	AGE_RATE_C	0.80	0.80	0.80	0.80	0.80	0.80	0.00	0.00
	AGE_RATE_N	0.66	0.77	0.80	0.83	0.96	0.80	0.04	0.05
	AGE_RATE_B	0.21	0.67	0.80	0.94	1.00	0.78	0.17	0.22
P2a	*DP_UL_C*	6.54	9.30	9.99	10.67	13.67	9.98	1.01	0.10
DP	*DP_UL_N*	5.11	7.37	7.98	8.59	12.13	7.99	0.90	0.11
impacted by UL	*DP_UL_B*	1.87	6.51	7.91	9.23	13.25	7.83	1.89	0.24
P2b	*NPPE_UL_C*	13.08	18.60	19.98	21.35	27.35	19.97	2.02	0.10
NPPE	*NPPE_UL_N*	4.56	16.78	19.81	22.94	41.05	19.94	4.55	0.23
impacted by UL or AGE	*NPPE_UL_B*	0.00	6.47	19.78	33.10	94.59	21.53	17.39	0.81
P2c	*NPPE_BOTH_C*	13.08	18.60	19.98	21.35	27.35	19.97	2.02	0.10
NPPE	*NPPE_BOTH_N*	8.92	30.58	35.61	40.70	68.84	35.73	7.46	0.21
impacted by UL & AGE	*NPPE_BOTH_B*	0.00	12.51	35.52	54.90	120.40	35.47	25.94	0.73

Note: (a) P2a is to test the impact of *UL* on *DP*; P2b tests the impact of *UL* or *AGE* on *NPPE*; P2c tests the impact of *UL* and *AGE* on *NPPE*. (b) Variable definition can be found in Appendix 4. (c) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs that are at least 99% efficient (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (d) This table continues Table 5 - 7.

Appendix 6 Test Two Simulated Efficiency Scores

(This table is an extension of Table 5 - 8)

Panel A: N = 24, iteration = 100											
	Variables	**MEAN**	**100%**	**99%**	**Range**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Base	*SCORE_GPPE*	0.8841	3.30	4.17	0.4933	0.5067	0.8195	0.9016	0.9695	0.0984	0.1113
P2a	*SCORE_DP_UL_C*	0.8841	3.30	4.17	0.4933	0.5067	0.8195	0.9016	0.9695	0.0984	0.1113
DP	*SCORE_DP_UL_N*	0.8777	2.95	3.56	0.4684	0.5316	0.8112	0.8951	0.9621	0.0996	0.1135
impacted by UL	*SCORE_DP_UL_B*	0.8457	2.77	3.17	0.5218	0.4782	0.7680	0.8568	0.9345	0.1123	0.1328
P2b	*SCORE_NPPE_UL_C*	0.8841	3.30	4.17	0.4933	0.5067	0.8195	0.9016	0.9695	0.0984	0.1113
NPPE	*SCORE_NPPE_UL_N*	0.8517	2.86	3.24	0.5020	0.4980	0.7765	0.8640	0.9414	0.1099	0.1290
impacted by UL or AGE	*SCORE_NPPE_UL_B*	0.8293	2.31	2.64	0.5527	0.4473	0.7485	0.8364	0.9216	0.1155	0.1392
P2c	*SCORE_NPPE_BOTH_C*	0.8841	3.30	4.17	0.4933	0.5067	0.8195	0.9016	0.9695	0.0984	0.1113
NPPE	*SCORE_NPPE_BOTH_N*	0.8533	2.83	3.20	0.5020	0.4980	0.7782	0.8663	0.9429	0.1091	0.1278
impacted by UL & AGE	*SCORE_NPPE_BOTH_B*	0.8290	2.28	2.61	0.5523	0.4477	0.7486	0.8360	0.9212	0.1154	0.1392
Panel B: N = 96, iteration = 100											
	Variables	**MEAN**	**100%**	**99%**	**Range**	**MIN**	**Q1**	**MED**	**Q3**	**SD**	**CV**
Base	*SCORE_GPPE*	0.8698	4.91	8.65	0.5295	0.4705	0.8063	0.8887	0.9493	0.0989	0.1138
P2a	*SCORE_DP_UL_C*	0.8698	4.91	8.65	0.5295	0.4705	0.8063	0.8887	0.9493	0.0989	0.1138
DP	*SCORE_DP_UL_N*	0.8526	3.92	5.58	0.5169	0.4831	0.7874	0.8690	0.9306	0.1002	0.1175
impacted by UL	*SCORE_DP_UL_B*	0.8026	3.46	4.11	0.5680	0.4320	0.7295	0.8095	0.8801	0.1085	0.1351
P2b	*SCORE_NPPE_UL_C*	0.8698	4.91	8.65	0.5295	0.4705	0.8063	0.8887	0.9493	0.0989	0.1138
NPPE	*SCORE_NPPE_UL_N*	0.8097	3.61	4.51	0.5849	0.4151	0.7377	0.8177	0.8880	0.1080	0.1334
impacted by UL or AGE	*SCORE_NPPE_UL_B*	0.7792	2.09	2.60	0.6131	0.3869	0.7048	0.7836	0.8597	0.1109	0.1423
P2c	*SCORE_NPPE_BOTH_C*	0.8698	4.91	8.65	0.5295	0.4705	0.8063	0.8887	0.9493	0.0989	0.1138
NPPE	*SCORE_NPPE_BOTH_N*	0.8122	3.58	4.57	0.5849	0.4151	0.7401	0.8205	0.8906	0.1074	0.1323
impacted by UL & AGE	*SCORE_NPPE_BOTH_B*	0.7791	2.06	2.57	0.6131	0.3869	0.7047	0.7835	0.8593	0.1108	0.1422

Note: (a) Variable definition can be found in Appendix 4. (b) The descriptive features include the mean value (MEAN), the number of efficient DMUs (100%), the number of DMUs that are at least 99% efficient (99%), the range of efficiency scores (RANGE), the minimal value (MIN), 25th percentile (Q1), 50th percentile (MED), 75th percentile (Q3), standard deviation (SD), and the coefficient of variation, which is the ratio of the standard deviation to the mean (CV). (c) All efficiency scores are generated from the CRS model. (d) This table continues Table 5 - 8.

Appendix 7 Test Two Variation in Financial DEA Results

(This table is an extension of Table 5 - 9)

Panel A: N = 24, iteration = 100									
	Constant			**Narrow**			**Broad**		
	Pearson	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**
P1	1.0000	1.0000	0.0000	0.9654	0.9357	0.0191	0.8163	0.7778	0.0566
DP impacted by UL	0.0000***	0.0000***		0.0000***	0.0000***		0.0005***	0.0007***	
P2a	1.0000	1.0000	0.0000	0.8346	0.7950	0.0498	0.7855	0.7492	0.0674
NPPE impacted by UL or AGE	0.0000***	0.0000***		0.0001***	0.0004***		0.0025***	0.0038***	
P2b	1.0000	1.0000	0.0000	0.8449	0.8068	0.0478	0.7863	0.7510	0.0675
NPPE impacted by UL & AGE	0.0000***	0.0000***		0.0001***	0.0003***		0.0025***	0.0038***	
Panel B: N = 96, iteration = 100									
	Constant			**Narrow**			**Broad**		
	Pearson	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**	**Pearson**	**Spearman's**	**MAD**
P1	1.0000	1.0000	0.0000	0.9724	0.9599	0.0232	0.8436	0.8181	0.0740
DP impacted by UL	0.0000***	0.0000***		0.0000***	0.0000***		0.0000***	0.0000***	
P2a	1.0000	1.0000	0.0000	0.8478	0.8208	0.0685	0.8256	0.7966	0.0933
NPPE impacted by UL or AGE	0.0000***	0.0000***		0.0000***	0.0000***		0.0000***	0.0000***	
P2b	1.0000	1.0000	0.0000	0.8558	0.8291	0.0661	0.8265	0.7977	0.0933
NPPE impacted by UL & AGE	0.0000***	0.0000***		0.0000***	0.0000***		0.0000***	0.0000***	

Note: (a) DEA models are the constant return of scale. (b) *** for significant level of < 0.01, ** for significant level of significant level of < 0.05, * for significant level of < 0.1. (3) the criteria for Financial DEA results are the Pearson Correlation (Pearson), the Spearman's Ranking Correlation (Spearman's) and the Mean Absolute Deviation (MAD). (c) The comparison point is the Financial DEA results generated using GPPE. (c) This table continues Table 5 - 9.

Appendix 8 Test Three Detailed Standard Industrial Classification

(From "detail for 48 industry portfolios" by Kenneth R. French) [41]

Panel A: Autos Automobiles and Trucks
2296-2296 Tire cord and fabric
2396-2396 Automotive trimmings, apparel findings & related products
3010-3011 Tires and inner tubes
3537-3537 Industrial trucks, tractors, trailers & stackers
3647-3647 Vehicular lighting equipment
3694-3694 Electrical equipment for internal combustion engines
3700-3700 Transportation equipment
3710-3710 Motor vehicles and motor vehicle equipment
3711-3711 Motor vehicles & passenger car bodies
3713-3713 Truck & bus bodies
3714-3714 Motor vehicle parts & accessories
3715-3715 Truck trailers
3716-3716 Motor homes
3792-3792 Travel trailers and campers
3790-3791 Miscellaneous transportation equipment
3799-3799 Miscellaneous transportation equipment
Panel B: Box Shipping Containers
2440-2449 Wood containers
2640-2659 Paperboard containers, box, drums, tubs
3220-3221 Glass containers
3410-3412 Metal cans and shipping containers
Panel C: Clothes Apparel
2300-2390 Apparel and other finished products
3020-3021 Rubber and plastics footwear
3100-3111 Leather tanning and finishing
3130-3131 Boot & shoe cut stock & findings
3140-3149 Footwear, except rubber
3150-3151 Leather gloves and mittens
3963-3965 Fasteners, buttons, needles, pins
Panel D: Food Products
2000-2009 Food and kindred products
2010-2019 Meat products
2020-2029 Dairy products
2030-2039 Canned & preserved fruits & vegetables
2040-2046 Flour and other grain mill products
2050-2059 Bakery products
2060-2063 Sugar and confectionery products
2070-2079 Fats and oils
2090-2092 Miscellaneous food preparations and kindred products
2095-2095 Roasted coffee
2098-2099 Miscellaneous food preparations

[41] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_48_ind_port.html

3995-3995	Burial caskets
Panel E: Gold Precious Metals	
1040-1049	Gold & silver ores
Panel F: Personal services	
7020-7021	Rooming and boarding houses
7030-7033	Camps and recreational vehicle parks
7200-7200	Services - personal
7210-7212	Services - laundry, cleaning & garment services
7214-7214	Services - diaper service
7215-7216	Services - coin-operated cleaners, dry cleaners
7217-7217	Services - carpet & upholstery cleaning
7219-7219	Services - Miscellaneous laundry & garment services
7220-7221	Services - photographic studios, portrait
7230-7231	Services - beauty shops
7240-7241	Services - barber shops
7250-7251	Services - shoe repair shops & shoeshine parlours
7260-7269	Services - funeral service & crematories
7270-7290	Services - Miscellaneous
7291-7291	Services - tax return
7292-7299	Services - Miscellaneous
7395-7395	Services - photofinishing labs (School pictures)
7500-7500	Services - auto repair, services & parking
7520-7529	Services - automobile parking
7530-7539	Services - automotive repair shops
7540-7549	Services - automotive services, except repair (car washes)
7600-7600	Services - Miscellaneous repair services
7620-7620	Services - Electrical repair shops
7622-7622	Services - Radio and TV repair shops
7623-7623	Services - Refrigeration and air conditioning service & repair
7629-7629	Services - Electrical & electronic repair shops
7630-7631	Services - Watch, clock, and jewellery repair
7640-7641	Services - Reupholster & furniture repair
7690-7699	Services - Miscellaneous repair shops & related services
8100-8199	Services - legal
8200-8299	Services - educational
8300-8399	Services - social services
8400-8499	Services - museums, art galleries, botanical and zoological
8600-8699	Services - membership organisations
8800-8899	Services - private households
7510-7515	Services - truck & auto rental and leasing

Appendix 9 Test Three Correlations between Variables

Panel A: Automobile Industry N = 253												
	AT	*ACT*	*INTAN*	*NPPE*	*COGS*	*XOPR*	*XSGA*	*LT*	*CEQ-PO*	*MKVALT*	*SALE*	*NI-PO*
AT	1.0000											
ACT	0.9854	1.0000										
INTAN	0.4384	0.4040	1.0000									
NPPE	0.9838	0.9497	0.4321	1.0000								
COGS	0.9718	0.9540	0.6065	0.9525	1.0000							
XOPR	0.9719	0.9566	0.6111	0.9521	0.9995	1.0000						
XSGA	0.9282	0.9354	0.6240	0.9047	0.9493	0.9588	1.0000					
LT	0.9974	0.9832	0.4142	0.9771	0.9631	0.9625	0.9132	1.0000				
CEQ-PO	0.9536	0.9390	0.5251	0.9511	0.9559	0.9590	0.9423	0.9298	1.0000			
MKVALT	0.7921	0.7659	0.3587	0.8315	0.7671	0.7748	0.8080	0.7737	0.8176	1.0000		
SALE	0.9740	0.9568	0.6048	0.9569	0.9994	0.9997	0.9563	0.9644	0.9618	0.7784	1.0000	
NI-PO	0.7501	0.7399	0.4517	0.7178	0.7797	0.7800	0.7474	0.7234	0.8240	0.6255	0.7795	1.0000
Panel B: Box Industry N = 53												
	AT	*ACT*	*INTAN*	*NPPE*	*COGS*	*XOPR*	*XSGA*	*LT*	*CEQ-PO*	*MKVALT*	*SALE*	*NI-PO*
AT	1.0000											
ACT	0.9434	1.0000										
INTAN	0.9811	0.9341	1.0000									
NPPE	0.9655	0.8622	0.9110	1.0000								
COGS	0.9554	0.9733	0.9368	0.8948	1.0000							
XOPR	0.9627	0.9648	0.9394	0.9133	0.9981	1.0000						
XSGA	0.8532	0.7109	0.7909	0.9138	0.7985	0.8337	1.0000					
LT	0.9310	0.9692	0.9345	0.8346	0.9450	0.9342	0.6653	1.0000				
CEQ-PO	0.8185	0.6284	0.7699	0.8899	0.6935	0.7274	0.9016	0.5531	1.0000			
MKVALT	0.4335	0.4572	0.4088	0.4273	0.4686	0.4727	0.4227	0.3487	0.4448	1.0000		
SALE	0.7569	0.8005	0.7327	0.7062	0.7812	0.7734	0.5617	0.6859	0.6518	0.4975	1.0000	
NI-PO	0.9621	0.9649	0.9347	0.9190	0.9968	0.9991	0.8368	0.9308	0.7316	0.4877	0.7782	1.0000

Panel C: Clothing Industry N = 160

	AT	*ACT*	*INTAN*	*NPPE*	*COGS*	*XOPR*	*XSGA*	*LT*	*CEQ-PO*	*MKVALT*	*SALE*	*NI-PO*
AT	1.0000											
ACT	0.9515	1.0000										
INTAN	0.5137	0.2333	1.0000									
NPPE	0.9181	0.9299	0.2492	1.0000								
COGS	0.9384	0.9828	0.2465	0.9002	1.0000							
XOPR	0.9612	0.9865	0.3049	0.9135	0.9944	1.0000						
XSGA	0.9752	0.9702	0.3884	0.9137	0.9637	0.9865	1.0000					
LT	0.9690	0.9160	0.5050	0.9089	0.8988	0.9181	0.9278	1.0000				
CEQ-PO	0.9380	0.9005	0.4722	0.8341	0.8948	0.9199	0.9385	0.8232	1.0000			
MKVALT	0.8803	0.9613	0.1074	0.9028	0.9572	0.9525	0.9242	0.8547	0.8231	1.0000		
SALE	0.9608	0.9855	0.3080	0.9087	0.9933	0.9993	0.9864	0.9128	0.9266	0.9507	1.0000	
NI-PO	0.8639	0.9170	0.2129	0.7975	0.9229	0.9206	0.8965	0.7787	0.8922	0.8886	0.9293	1.0000

Panel D: Food Industry N= 218

	AT	*ACT*	*INTAN*	*NPPE*	*COGS*	*XOPR*	*XSGA*	*LT*	*CEQ-PO*	*MKVALT*	*SALE*	*NI-PO*
AT	1.0000											
ACT	0.6513	1.0000										
INTAN	0.9505	0.3921	1.0000									
NPPE	0.7920	0.9136	0.5740	1.0000								
COGS	0.4912	0.9420	0.2228	0.8430	1.0000							
XOPR	0.5390	0.9524	0.2720	0.8810	0.9960	1.0000						
XSGA	0.7222	0.5767	0.6011	0.8016	0.4631	0.5402	1.0000					
LT	0.9865	0.6787	0.9166	0.8358	0.5235	0.5768	0.7938	1.0000				
CEQ-PO	0.9834	0.6022	0.9559	0.7249	0.4433	0.4841	0.6294	0.9420	1.0000			
MKVALT	0.9159	0.5622	0.8665	0.7643	0.4024	0.4659	0.8338	0.9288	0.8742	1.0000		
SALE	0.6054	0.9573	0.3489	0.9140	0.9857	0.9957	0.5957	0.6434	0.5490	0.5395	1.0000	
NI-PO	0.4407	0.3328	0.3748	0.4713	0.2961	0.3328	0.5140	0.4523	0.4225	0.5798	0.3700	1.0000

Panel E: Gold Industry N = 43

	AT	*ACT*	*INTAN*	*NPPE*	*COGS*	*XOPR*	*XSGA*	*LT*	*CEQ-PO*	*MKVALT*	*SALE*	*NI-PO*
AT	1.0000											

ACT	0.9564	1.0000										
INTAN	0.8776	0.7908	1.0000									
NPPE	0.9867	0.9080	0.8572	1.0000								
COGS	0.8623	0.8731	0.6052	0.8603	1.0000							
XOPR	0.8784	0.8900	0.6254	0.8746	0.9988	1.0000						
XSGA	0.9283	0.9460	0.7313	0.9103	0.8984	0.9188	1.0000					
LT	0.9560	0.9494	0.7701	0.9368	0.9190	0.9326	0.9568	1.0000				
CEQ-PO	0.9544	0.8702	0.8724	0.9660	0.7578	0.7752	0.8448	0.8389	1.0000			
MKVALT	0.8876	0.8178	0.8662	0.8857	0.6326	0.6596	0.8177	0.7972	0.9142	1.0000		
SALE	0.9609	0.9814	0.7689	0.9289	0.9081	0.9242	0.9696	0.9797	0.8610	0.8190	1.0000	
NI-PO	0.3874	0.3344	0.5423	0.3591	-	-	0.2258	0.1952	0.5239	0.5947	0.2750	1.0000
Panel F: Personal Services Industry N= 47												
	AT	*ACT*	*INTAN*	*NPPE*	*COGS*	*XOPR*	*XSGA*	*LT*	*CEQ-PO*	*MKVALT*	*SALE*	*NI-PO*
AT	*1.0000*											
ACT	0.3629	1.0000										
INTAN	0.8134	0.6125	1.0000									
NPPE	0.9600	0.3709	0.8352	1.0000								
COGS	0.7204	0.6885	0.8741	0.7291	1.0000							
XOPR	0.6212	0.7920	0.8852	0.6590	0.9680	1.0000						
XSGA	0.3202	0.8445	0.7455	0.4062	0.7313	0.8792	1.0000					
LT	0.9738	0.2684	0.7184	0.8923	0.5990	0.4856	0.1823	1.0000				
CEQ-PO	0.6349	0.5186	0.7663	0.7536	0.8057	0.7996	0.6428	0.4424	1.0000			
MKVALT	0.6049	0.7307	0.8941	0.6745	0.8423	0.9211	0.9031	0.4594	0.8240	1.0000		
SALE	0.6551	0.7998	0.9008	0.6831	0.9600	0.9951	0.8810	0.5281	0.7886	0.9333	1.0000	
NI-PO	0.6214	0.7281	0.8218	0.6405	0.8201	0.8814	0.8378	0.5091	0.7206	0.9145	0.9151	1.0000

Note: (a) The numbers in bold have p-values greater than 0.01. (b) The numbers in grey shadows are less than 0.2, showing a weak correlation. (c) *AT*, total assets. *ACT*, current assets. *NPPE*, net Property, Plant and Equipment. INTAN, intangibles. XOPR, operating expenses. *XSGA*, selling, general and administrative expenses. *COGS*, cost of goods sold. *LT*, total liabilities. *NI-PO*, net income, translated to a positive value. *CEQ-PO*, common equity, translated to a positive value. *MKVALT*, market value. *SALE*, sales.

Appendix 10 Test Three B Variation in Financial DEA Results

(This table reports the Pearson correlations, which is an extension of Table 5 - 22)

Panel A: Automobile Industry N = 253								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	-0.2319	1.0000						
Model III	0.2962	0.3916	1.0000					
Model IV	0.9926	-0.2666	0.2776	1.0000				
Model V	-0.2152	0.9486	0.3901	-0.2194	1.0000			
Model VI	0.3196	0.2865	0.9653	0.3311	0.3558	1.0000		
Model VII	-0.2279	0.9142	0.3135	-0.2365	0.9598	0.2710	1.0000	
Model VIII	0.0148	-0.1536	0.2578	0.0087	-0.1795	0.2642	-0.2537	1.0000
Panel B: Box Industry N = 53								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.4166	1.0000						
Model III	0.9683	0.4825	1.0000					
Model IV	0.8256	0.3058	0.8407	1.0000				
Model V	0.5349	0.7861	0.5939	0.4262	1.0000			
Model VI	0.8181	0.3131	0.8442	0.9946	0.4301	1.0000		
Model VII	0.5357	0.7974	0.5816	0.3966	0.9607	0.3956	1.0000	
Model VIII	0.3292	0.1843	0.2980	0.1348	0.0429	0.1249	0.0706	1.0000
Panel C: Clothing Industry N = 160								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.0565	1.0000						
Model III	0.6190	0.5358	1.0000					
Model IV	0.9062	-0.0895	0.4506	1.0000				
Model V	0.0432	0.8589	0.4714	0.0458	1.0000			

Model VI	0.5671	0.3345	0.8639	0.6062	0.4789	1.0000		
Model VII	-0.0957	0.7578	0.3390	-0.0535	0.9256	0.3878	1.0000	
Model VIII	0.2704	-0.1446	0.2090	0.2686	-0.0825	0.2813	-0.1242	1.0000
Panel D: Food Industry N = 218								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	-0.4141	1.0000						
Model III	0.0660	0.2785	1.0000					
Model IV	0.9672	-0.4025	0.0762	1.0000				
Model V	-0.2345	0.8340	0.2758	-0.2210	1.0000			
Model VI	0.0903	0.2546	0.9755	0.1122	0.2870	1.0000		
Model VII	-0.2172	0.7848	0.3140	-0.2045	0.9761	0.3225	1.0000	
Model VIII	0.1226	-0.1599	0.0051	0.1100	-0.2443	0.0059	-0.2843	1.0000
Panel E: Gold Industry N = 43								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.2498	1.0000						
Model III	0.7358	0.5508	1.0000					
Model IV	0.9836	0.2722	0.7503	1.0000				
Model V	0.2750	0.9943	0.5725	0.2965	1.0000			
Model VI	0.7354	0.5493	0.9999	0.7505	0.5709	1.0000		
Model VII	0.2588	0.9901	0.5657	0.2814	0.9966	0.5641	1.0000	
Model VIII	-0.0769	-0.4421	-0.2117	-0.1106	-0.4243	-0.2166	-0.4251	1.0000
Panel F: Personal Services Industry N = 47								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.1262	1.0000						
Model III	0.4831	0.6440	1.0000					
Model IV	0.9828	0.1015	0.4539	1.0000				
Model V	0.1225	0.9143	0.5786	0.1731	1.0000			
Model VI	0.4717	0.5955	0.9484	0.5021	0.6611	1.0000		

Model VII	-0.0080	0.7750	0.4233	0.0505	0.8804	0.5181	1.0000	
Model VIII	-0.3474	0.3010	-0.0248	-0.3666	0.2297	-0.0670	0.2603	1.0000

Note: (a) DEA models are the variable return of scale output-orientation. The variable return of scale, input-orientation models, generate similar results, untabulated. (b) Model I was adapted from Demerjian et al. (2012) Demerjian et al. (2013and Demerjian (2018). Model II was the transit model between Model I and Model III. Model III was adapted from Harrison and Rouse (2016). Model IV was the DuPont ratio model adapted from Feroz et al. (2001,2003). Model V was the transit model between Model IV and VI. Model VI was adapted from the stage one DEA model from Seiford and Zhu (1999) and Zhu (2000). Model VII was the funding model adapted from Smith (1990). Model VIII was adapted from the stage two DEA model from Seiford and Zhu (1999) and Zhu (2000). (c) The correlation coefficients in bold demonstrate discriminant validity. The cut-off of the discriminant validity is the correlation coefficient < 0.2 or p-value > 0.1. (d) The Spearman's correlation results can be found in Table 5 - 22.

Appendix 11 Test Three B with Subsamples Variation in Financial DEA Results

(This table reports the Pearson correlations, which is an extension of Table 5 - 24)

Panel A: Automobile Industry N = 253								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.5634	1.0000						
Model III	0.9639	0.6468	1.0000					
Model IV	0.9736	0.5180	0.9355	1.0000				
Model V	0.5049	0.9024	0.5834	0.5345	1.0000			
Model VI	0.9213	0.6063	0.9588	0.9560	0.6351	1.0000		
Model VII	0.5222	0.8445	0.5938	0.5376	0.9262	0.6288	0.1661	
Model VIII	0.2696	0.2864	0.2538	0.2167	0.2070	0.1861	1.0000	1.0000
Panel B: Box Industry N = 53								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.6390	1.0000						
Model III	1.0000	0.6390	1.0000					
Model IV	0.7420	0.3906	0.7420	1.0000				
Model V	0.5516	0.7983	0.5516	0.6040	1.0000			
Model VI	0.7420	0.3906	0.7420	1.0000	0.6040	1.0000		
Model VII	0.5926	0.8256	0.5926	0.5789	0.9866	0.5789	0.1936	
Model VIII	0.3289	0.3225	0.3289	-0.0066	0.1445	-0.0066	1.0000	1.0000
Panel C: Clothing Industry N = 160								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.6102	1.0000						
Model III	0.9550	0.6790	1.0000					
Model IV	0.9148	0.5395	0.8886	1.0000				
Model V	0.6045	0.8737	0.6598	0.6582	1.0000			

Model VI	0.8755	0.5860	0.9152	0.9757	0.6973	1.0000		
Model VII	0.4540	0.6758	0.5113	0.5190	0.8437	0.5606	0.2443	
Model VIII	0.3577	0.2614	0.3240	0.2994	0.2455	0.2699	1.0000	1.0000
Panel D: Food Industry N = 218								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.4760	1.0000						
Model III	0.9948	0.5045	1.0000					
Model IV	0.9478	0.4274	0.9416	1.0000				
Model V	0.4520	0.9315	0.4776	0.4933	1.0000			
Model VI	0.9429	0.4607	0.9478	0.9934	0.5267	1.0000		
Model VII	0.4420	0.8793	0.4642	0.4806	0.9378	0.5061	0.0452	
Model VIII	0.0010	0.0574	0.0029	0.0496	0.1032	0.0502	1.0000	1.0000
Panel E: Gold Industry N = 43								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.6282	1.0000						
Model III	1.0000	0.6285	1.0000					
Model IV	1.0000	0.6282	1.0000	1.0000				
Model V	0.6618	0.9843	0.6621	0.6618	1.0000			
Model VI	1.0000	0.6285	1.0000	1.0000	0.6621	1.0000		
Model VII	0.6296	0.9340	0.6299	0.6296	0.9505	0.6299	-0.1453	
Model VIII	-0.1378	-0.1771	-0.1388	-0.1378	-0.1395	-0.1388	1.0000	1.0000
Panel F: Personal Services Industry N = 47								
	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*	*Model VII*	*Model VIII*
Model I	1.0000							
Model II	0.5525	1.0000						
Model III	0.9631	0.6124	1.0000					
Model IV	1.0000	0.5535	0.9630	1.0000				
Model V	0.6486	0.9236	0.7000	0.6493	1.0000			
Model VI	0.9631	0.6124	1.0000	0.9630	0.7002	1.0000		

Model VII	0.7156	0.7183	0.7459	0.7164	0.7987	0.7462	-0.0453	
Model VIII	0.1821	0.0347	0.2002	0.1817	-0.1056	0.1989	1.0000	1.0000

Note: (a) The numbers in bold have p-values greater than 0.01. (b) The numbers in grey shadows are less than 0.2, which demonstrate discriminant validity. (c) DEA models are the variable return of scale output-orientation. The variable return of scale, input-orientation models, generate similar results, untabulated. (d) Model I was adapted from Demerjian et al. (2012), Demerjian et al. (2013) and Demerjian (2018). Model II was the transit model between Model I and Model III. Model III was adapted from Harrison and Rouse (2016). Model IV was the DuPont ratio model adapted from Feroz et al. (2001,2003). Model V was the transit model between Model IV and VI. Model VI was adapted from the stage one DEA model from Seiford and Zhu (1999) and Zhu (2000). Model VII was the funding model adapted from Smith (1990). Model VIII was adapted from the stage two DEA model from Seiford and Zhu (1999) and Zhu (2000). (e) The Spearman's ranking correlation results can be found in Table 5 - 24.

List of References

Ahn, H., Choi, S., & Yun, S. C. (2020). Financial statement comparability and the market value of cash holdings. *Accounting Horizons, 34*(3), 1-21.

Akdeniz, M. B., Gonzalez-Padron, T., & Calantone, R. J. (2010). An integrated marketing capability benchmarking approach to dealer performance through parametric and nonparametric analyses. *Industrial Marketing Management, 39*(1), 150-160.

Al-Sharkas, A. A., Hassan, M. K., & Lawrence, S. (2008). The impact of mergers and acquisitions on the efficiency of the US banking industry: Further evidence. *Journal of Business Finance and Accounting, 35*(1-2), 50-70.

Ali, I., & Seiford, L. (1990). Translation invariance in Data Envelopment Analysis. *Operations Research Letters, 9*(6), 403-405.

Amado, C. A. F., Santos, S. P., & Marques, P. M. (2012). Integrating the Data Envelopment Analysis and the Balanced Scorecard approaches for enhanced performance assessment. *Omega, 40*(3), 390-403.

American Psychological Association. (2020). *Dictionary of Psychology.*

Andon, P., Baxter, J., & Chua, W. F. (2015). Accounting for stakeholders and making accounting useful. *Journal of Management Studies, 52*(7), 986-1002.

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). *Journal of Productivity Analysis, 41*(1), 85-109.

Andor, M., & Hesse, F. (2011). A Monte Carlo Simulation comparing DEA, SFA and two simple approaches to combine efficiency estimates. Paper presented at the *CAWM Discussion Paper/Centrum Für Angewandte Wirtschaftsforschung Münster.*

Andreou, P. C., Karasamani, I., Louca, C., & Ehrlich, D. (2017). The impact of managerial ability on crisis-period corporate investment. *Journal of Business Research, 79*, 107-122.

Anglo Gold Ashanti. (2019). *Financial Report of Anglo Gold Ashanti LTD 2019.*

Aparicio, J., & Kapelko, M. (2018). Accounting for slacks to measure dynamic inefficiency in Data Envelopment Analysis. *European Journal of Operational Research. 278*(2), 463-471.

Arbnor, I., & Bjerke, B. (2009). *Methodology for Creating Business Knowledge* (3rd ed.). Sage.

Asmild, M., & Tam, F. (2007). Estimating global frontier shifts and global Malmquist indices. *Journal of Productivity Analysis, 27*(2), 137-148.

Athanassopoulos, A. D. (1995). Goal programming & Data Envelopment Analysis (GoDEA) for target-based multi-level planning: Allocating central grants to the Greek local authorities. *European Journal of Operational Research, 87*(3), 535-550.

Athanassopoulos, A. D. (1997). Service quality and operating efficiency synergies for management control in the provision of financial services: Evidence from Greek bank branches. *European Journal of Operational Research, 98*(2), 300-313.

Avilés-Sacoto, S. V., Cook, W. D., & Güemes-Castorena, D. (2014). Competitiveness among higher education institutions: A two-stage Cobb-Douglas model for efficiency measurement of Schools of Business. *Journal of Centrum Cathedra: The Business and Economics Research Journal, 7*(1), 91-115.

Avkiran, N. K. (1999). The evidence on efficiency gains: The role of mergers and the benefits to the public. *Journal of Banking and Finance, 23*(7), 991-1013.

Avkiran, N. K. (2000). Rising productivity of Australian trading banks under deregulation 1986–1995. *Journal of Economics and Finance, 24*(2), 122-140.

Avkiran, N. K. (2009). Removing the impact of environment with units-invariant efficient frontier analysis: An illustrative case study with intertemporal panel data. *Omega, 37*(3), 535-544.

Avkiran, N. K. (2011). Association of DEA super-efficiency estimates with financial ratios: Investigating the case for Chinese banks. *Omega, 39*(3), 323-334.

Babbie, E. R. (2016). *The Practice of Social Research* (14th ed.). Cengage Learning.

Baghdadi, G. A., Bhatti, I. M., Nguyen, L. H. G., & Podolski, E. J. (2018). Skill or effort? Institutional ownership and managerial efficiency. *Journal of Banking and Finance, 91*, 19-33.

Baik, B., Chae, J., Choi, S., & Farber, D. B. (2013). Changes in operational efficiency and firm performance: A frontier analysis approach. *Contemporary Accounting Research, 30*(3), 996-1026.

Baik, B., Farber, D. B., & Lee, S. (2011). CEO ability and management earnings forecasts. *Contemporary Accounting Research, 28*(5), 1645-1668.

Bailey, K. D. (1994). *Typologies and Taxonomies an Introduction to Classification Techniques*. Sage.

Banker, R. D., & Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research, 175*(2), 1311-1320.

Banker, R. D., Chang, H., & Kao, Y. (2002). Impact of information technology on public accounting firm productivity. *Journal of Information Systems, 16*(2), 209.

Banker, R. D., Chang, H., & Natarajan, R. (2007). Estimating DEA technical and allocative inefficiency using aggregate cost or revenue data. *Journal of Productivity Analysis, 27*(2), 115-121.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management science, 30*(9), 1078-1092.

Banker, R. D., Charnes, A., Cooper, W. W., Swarts, J., & Thomas, D. A. (1989). An introduction to Data Envelopment Analysis with some of its models and their uses. *Research in Governmental and Non-profit Accounting, 5*(5), 125-163.

Banker, R. D., Darrough, M. N., Huang, R., & Plehn-Dujowich, J. M. (2013). The relation between CEO compensation and past performance. *The Accounting Review, 88*(1), 1-30.

Banker, R. D., Fang, S., & Mehta, M. N. (2019). Anomalous operating performance during economic slowdowns. *Journal of Management Accounting Research, 32*(2), 57.

Banker, R. D., Gadh, V. M., & Gorr, W. L. (1993). A Monte Carlo comparison of two production frontier estimation methods: Corrected ordinary least squares and Data Envelopment Analysis. *European Journal of Operational Research, 67*(3), 332-343.

Banker, R. D., Huang, R., Li, Y., & Zhao, S. (2021). Do accounting standards matter for productivity? *Production and Operations Management, 1*(30).

Bao, M. X., Cheng, X., Smith, D., & Tanyi, P. (2021). CEO pay ratios and financial reporting quality. *Global Finance Journal, 47*, 100506.

Bardhan, I., Bowlin, W. F., Cooper, W. W., & Sueyoshi, T. (1996). Models and measures for efficiency dominance in DEA Part 1: Additive models and MED measures. *Journal of the Operations Research Society of Japan, 39*(3), 322-332.

Barth, M. E., Landsman, W. R., Lang, M., & Williams, C. (2012). Are IFRS-based and US GAAP-based accounting amounts comparable? *Journal of Accounting and Economics, 54*(1), 68-93.

Barth, M. E., Li, K., & McClure, C. (2021). Evolution in value relevance of accounting information. Working paper.

Belton, V., & Vickers, S. P. (1993). Demystifying DEA—A visual interactive approach based on multiple criteria analysis. *Journal of the Operational Research Society, 44*(9), 883-896.

Benli, Y. K., & Bozoklu, Ç P. (2016). Complexity of measuring advertising efficiency: An application of DEA method in Turkey. *Chaos, Complexity and Leadership*, 193-207. Springer.

Bennouri, M., Chtioui, T., Nagati, H., & Nekhili, M. (2018). Female board directorship and firm performance: What really matters? *Journal of Banking and Finance, 88*, 267-291.

Benston, G. J. (2004). What's special about banks? *Financial Review, 39*(1), 13-33.

Berger, A. N., Hasan, I., & Zhou, M. (2009). Bank ownership and efficiency in China: What will happen in the world's largest nation? *Journal of Banking and Finance, 33*(1), 113-130.

Bhattacharyya, A., Lovell, C. A. K., & Sahay, P. (1997). The impact of liberalization on the productive efficiency of Indian commercial banks. *European Journal of Operational Research, 98*(2), 332-345.

Bisbe, J., Batista-Foguet, J., & Chenhall, R. (2007). Defining management accounting constructs: A methodological note on the risks of conceptual misspecification. *Accounting, Organizations and Society, 32*(7), 789-820.

Blackburn, S. (2016). *The Oxford dictionary of philosophy* (3rd ed.). Oxford University Press.

Blalock, H. M. (1971). Causal models involving unobserved variables in stimulus-response situations. *Causal Models in the Social Sciences*, 335-347. Aldine.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons.

Bollen, K. A., & Ting, K. (2000). A Tetrad Test for Causal Indicators. *Psychological Methods, 5*(1), 3-22.

Bonsall IV, S. B., Holzman, E. R., & Miller, B. P. (2017). Managerial ability and credit risk assessment. *Management Science, 63*(5), 1425-1449.

Boussofiane, A., Dyson, R. G., & Thanassoulis, E. (1991). Applied Data Envelopment Analysis. *European Journal of Operational Research, 52*(1), 1-15.

Bowlin, W. F. (1995). A characterisation of the financial condition of the United States' aerospace-defense industrial base. *Omega, 23*(5), 539-555.

Bowlin, W. F. (1998). Measuring performance: An introduction to Data Envelopment Analysis (DEA). *The Journal of Cost Analysis, 15*(2), 3-27.

Bowlin, W. F. (1999). An analysis of the financial performance of defense business segments using Data Envelopment Analysis. *Journal of Accounting and Public Policy, 18*(4), 287-310.

Bowlin, W. F., & Renner, C. J. (2008). Assessing gender and top-management-team pay in the S & P Mid-Cap and Small-Cap companies using Data Envelopment Analysis. *European Journal of Operational Research, 185*(1), 430.

Brockett, P. L., Charnes, A., Cooper, W. W., Huang, Z. M., & Sun, D. B. (1997). Data transformations in DEA cone ratio envelopment approaches for monitoring bank performances. *European Journal of Operational Research, 98*(2), 250-268.

Brooks, M., Hairston, S., & Harter, C. (2020). Does manager ability influence the classification of lease arrangements? *Journal of Applied Accounting Research, 21*(1), 19-37.

Bryan, D., Fernando, G. D., & Tripathy, A. (2013). Bankruptcy risk, productivity and firm strategy. *Review of Accounting and Finance, 12*(4), 309-326.

Bui, D. G., Chen, Y., Hasan, I., & Lin, C. (2018). Can lenders discern managerial ability from luck? Evidence from bank loan contracts. *Journal of Banking and Finance, 87*, 187-201.

Burrell, G., & Morgan, G. (1979). In Morgan G. (Ed.), *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life*. Heinemann.

Camanho, A. S., & Dyson, R. G. (2005). Cost efficiency measurement with price uncertainty: a DEA application to bank branch assessments. *European Journal of Operational Research, 161*(2), 432-446.

Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist, 15*(8), 546.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81.

Cappa, F., Oriani, R., Peruffo, E., & McCarthy, I. (2021). Big data for creating and capturing value in the digitalised environment: unpacking the effects of volume, variety, and veracity on firm performance. *Journal of Product Innovation Management, 38*(1), 49-67.

Chang, H., Fernando, G. D., & Tripathy, A. (2015). An Empirical Study of Strategic Positioning and Production Efficiency. *Advances in Operations Research, 2015*(1), 1-11.

Chang, H., Fernando, G., Srinivasan, D., & Tripathy, A. (2013). A Re-Examination of Diversification and Firm Productivity. *Journal of Management Accounting Research, 25*(1), 99-118.

Chang, W., Hayes, R. M., & Hillegeist, S. A. (2016). Financial distress risk and new CEO compensation. *Management Science, 62*(2), 479-501.

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. *European Journal of Operational Research, 279*(3), 929-940.

Charnes, A., Cooper, W. W., Huang, Z. M., & Sun, D. B. (1990). Polyhedral Cone-Ratio DEA Models with an illustrative application to large commercial banks. *Journal of Econometrics, 46*(1), 73-91.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research, 2*(6), 429-444.

Charnes, A., Cooper, W. W., Seiford, L., & Stutz, J. (1983). Invariant multiplicative efficiency and piecewise Cobb-Douglas envelopments. *Operations Research Letters, 2*(3), 101-103.

Charnes, A., Cooper, W. W., & Thrall, R. M. (1986). Classifying and characterizing efficiencies and inefficiencies in Data Envelopment Analysis. *Operations Research Letters, 5*(3), 105-110.

Chen, F., Liu, Z., & Kweh, Q. L. (2014). Intellectual capital and productivity of Malaysian general insurers. *Economic Modelling, 36*, 413-420.

Chen, J., & Chen, J. (2020). Does managerial ability affect the quality of environmental financial disclosure? *Sustainability Accounting, Management and Policy Journal, 11*(6), 1055-1073.

Chen, M., Ruan, L., Zhu, Z., & Sang, F. (2020). Macro uncertainty, analyst performance, and managerial ability. *Eurasian Business Review, 10*(3), 333-353.

Chen, T., Chen, C., & Peng, S. (2008). Firm operation performance analysis using Data Envelopment Analysis and balanced scorecard. *International Journal of Productivity and Performance Management,*

Chen, Y., & Chu, H. (2020). Empirical research on the function of compensation committees: the influence of family control and professional managers. *Asia-Pacific Journal of Accounting and Economics, 27*(1), 51-70.

Chen, Y., Podolski, E. J., & Veeraraghavan, M. (2015). Does managerial ability facilitate corporate innovative success? *Journal of Empirical Finance, 34,* 313-326.

Chen, Y., & Chen, B. (2011). Applying DEA, MPI, and grey model to explore the operation performance of the Taiwanese wafer fabrication industry. *Technological Forecasting and Social Change, 78*(3), 536-546.

Cheng, Q., Goh, B. W., & Kim, J. B. (2018). Internal control and operational efficiency. *Contemporary Accounting Research, 35*(2), 1102-1139.

Cheung, K. T. S., Naidu, D., Navissi, F., & Ranjeeni, K. (2017). Valuing talent: Do CEOs' ability and discretion unambiguously increase firm performance. *Journal of Corporate Finance, 42,* 15-35.

Choo, M., Wang, C., Yin, C., & Li, J. (2021). Managerial ability and external financing. *Asia-Pacific Financial Markets, 28*(2), 207-241.

Chua, W. F. (1986). Radical developments in accounting thought. *The Accounting Review, 61*(4), 601-632.

Clarke, V., & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology, 12*(3), 297-298.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). Springer.

Combs, J. G., Crook, T. R., & Shook, C. L. (2005). The dimensionality of organizational performance and its implications for strategic management research. *Research Methodology in Strategy and Management, 2*(5), 259-286.

Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA) – Thirty years on. *European Journal of Operational Research, 192*(1), 1-17.

Cook, W. D., Tone, K., & Zhu, J. (2014). Data Envelopment Analysis: Prior to choosing a model. *Omega, 44*, 1-4.

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data Envelopment Analysis a comprehensive text with models, applications, references and DEA-solver software* (2nd ed.). SpringerLink.

Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). Handbook on Data Envelopment Analysis.

Cordero, J. M., Pedraja, F., & Santín, D. (2009). Alternative approaches to include exogenous variables in DEA measures: A comparison using Monte Carlo. *Computers and Operations Research, 36*(10), 2699-2706.

Cornaggia, K. J., Krishnan, G. V., & Wang, C. (2017). Managerial ability and credit ratings. *Contemporary Accounting Research, 34*(4), 2094-2122.

Creamer, E. G. (2018). *An introduction to fully integrated mixed methods research*. Sage.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302.

Cross, R. M., & Färe, R. (2008). Farrell efficiency under value and quantity data. *Journal of Productivity Analysis, 29*(3), 193-199.

Cross, R. M., & Färe, R. (2009). Value data and the Bennet price and quantity indicators. *Economics Letters, 102*(1), 19-21.

Cui, H., & Chi-Moon Leung, S. (2020). The long-run performance of acquiring firms in mergers and acquisitions: Does managerial ability matter? *Journal of Contemporary Accounting and Economics, 16*(1), 100185.

da Silva, A. F., Marins, F. A. S., Dias, E. X., & Ushizima, C. A. (2021). Improving manufacturing cycle efficiency through new multiple criteria Data Envelopment Analysis models: an application in green and lean manufacturing processes. *Production Planning and Control, 32*(2), 104-120.

Debreu, G. (1951). The Coefficient of Resource Utilization. *Econometrica, 19*(3), 273-292.

Delmas, M., Russo, M. V., & Montes-Sancho, M. (2007). Deregulation and environmental differentiation in the electric utility industry. *Strategic Management Journal, 28*(2), 189-209.

Demerjian, P. R. (2018). Calculating efficiency with financial accounting data: Data Envelopment Analysis for accounting researchers. Working Paper.

Demerjian, P. R., Lev, B., Lewis, M. F., & McVay, S. E. (2013). Managerial ability and earnings quality. *Accounting Review, 88*(2), 463-498.

Demerjian, P. R., Lev, B., & McVay, S. (2012). Quantifying managerial ability: A new measure and validity tests. *Management Science, 58*(7), 1229-1248.

Demerjian, P. R., Lewis-Western, M., & McVay, S. (2020). How does intentional earnings smoothing vary with managerial ability? *Journal of Accounting, Auditing AND Finance, 35*(2), 406-437.

Diewert, W. E., & Mendoza, M. N. F. (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.

Doyle, J., & Green, R. (1993). Data Envelopment Analysis and multiple criteria decision making. *Omega, 21*(6), 713-715.

Durand, R., & Vargas, V. (2003). Ownership, organization, and private firms' efficient use of resources. *Strategic Management Journal, 24*(7), 667-675.

Dutta, A. (2013). Impact of privatization on productivity: A non-parametric analysis of Indian insurance sector. *Global Business Review, 14*(2), 297-314.

Dutta, A., & Sengupta, P. P. (2011). Efficiency measurement of Indian life insurance industry in post-reforms era. *Global Business Review, 12*(3), 415-430.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research, 132*(2), 245-259.

Edirisinghe, N. C. P., & Zhang, X. (2008). Portfolio selection under DEA-based relative financial strength indicators: case of US industries. *Journal of the Operational Research Society, 59*(6), 842-856.

Edwards, J. R. (2003). Construct validation in organizational behavior research. (pp. 327-371). Lawrence Erlbaum Associates Publishers.

Edwards, J. R., & Bagozzi, R. P. (2000). On the Nature and Direction of Relationships Between Constructs and Measures. *Psychological Methods, 5*(2), 155-174.

Efron, B. (1993). In Tibshirani R. (Ed.), *An introduction to the bootstrap*. Chapman & Hall.

Evans, J. H., Luo, S., & Nagarajan, N. J. (2014). CEO turnover, financial distress, and contractual innovations. *The Accounting Review, 89*(3), 959-990.

Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics, 43*(2), 153-193.

Fanchon, P. (2003). Variable selection for dynamic measures of efficiency in the computer industry. *International Advances in Economic Research, 9*(3), 175-188.

Färe, R., & Grosskopf, S. (1985). A nonparametric cost approach to scale efficiency. *The Scandinavian Journal of Economics*, 594-604.

Färe, R., Grosskopf, S., Karagiannis, G., & Margaritis, D. (2017). Data Envelopment Analysis and its related linear programming models. *Annals of Operations Research, 250*(1), 37-43.

Färe, R., Grosskopf, S., & Lee, H. (1990). A nonparametric approach to expenditure-constrained profit maximization. *American Journal of Agricultural Economics, 72*(3), 574-581.

Färe, R., Grosskopf, S., Lundström, M., & Roos, P. (2007). Evaluating health care efficiency. *Evaluating Hospital Policy and Performance: Contributions from Hospital Policy and Productivity Research*. Emerald Group Publishing Limited.

Färe, R., Grosskopf, S., & Weber, W. L. (2006). Shadow prices and pollution costs in US agriculture. *Ecological Economics, 56*(1), 89-103.

Färe, R., Grosskopf, S., & Zelenyuk, V. (2004). Aggregation bias and its bounds in measuring technical efficiency. *Applied Economics Letters, 11*(10), 657-660.

Färe, R., & Primont, D. (1995). *Multi-output production and duality: theory and applications*. Kluwer Academic Publishers.

Färe, R., & Zelenyuk, V. (2002). Input aggregation and technical efficiency. *Applied Economics Letters, 9*(10), 635-636.

Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). *The Measurement of Efficiency of Production.* (Vol. 6). Springer Science & Business Media.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, 120*(3), 253-290.

Fernandes, F. D. S., Stasinakis, C., & Bardarova, V. (2018). Two-stage DEA-Truncated Regression: Application in banking efficiency and financial development. *Expert Systems with Applications, 96*, 284-301.

Fernando, G. D., Jain, S. S., & Tripathy, A. (2020). This cloud has a silver lining: Gender diversity, managerial ability, and firm performance. *Journal of Business Research, 117*, 484-496.

Feroz, E. H., Kim, S., & Raab, R. L. (2003). Financial statement analysis: A Data Envelopment Analysis approach. *Journal of the Operational Research Society, 54*(1), 48.

Feroz, E. H., Raab, R., & Haag, S. (2001). An income efficiency model approach to the economic Consequences of OSHA cotton dust regulation. *Australian Journal of Management, 26*(1), 69-89.

Fischer, H. (2010). *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer Science & Business Media.

Flew, A. (2002). *A Dictionary of Philosophy*. Pan 2002.

Francis, B., Hasan, I., Mani, S., & Ye, P. (2016). Relative peer quality and firm performance. *Journal of Financial Economics, 122*(1), 196-219.

Francis, B., Hasan, I., Siraj, I., & Wu, Q. (2020). Managerial ability and value relevance of earnings. *China Accounting and Finance Review, Forthcoming,*

Frijns, B., Margaritis, D., & Psillaki, M. (2012). Firm efficiency and stock returns. *Journal of Productivity Analysis, 37*(3), 295-306.

Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review, 84*(1), 66-83.

Gaganis, C., Liadaki, A., Doumpos, M., & Zopounidis, C. (2009). Estimating and analyzing the efficiency and productivity of bank branches. *Managerial Finance, 35*(2), 202-218.

Gao, K., Wang, C., Su, Q., & Shen, H. (2020). Do analysts follow firms with able managers? *Managerial and Decision Economics, 41*(8), 1602-1612.

Gao, L., Rezaee, Z., & Yu, J. (2020). Peer firms' earnings predictability and pricing efficiency - evidence from IPOs. *The European Journal of Finance, 26*(13), 1332-1353.

Garrett, J., Hoitash, R., & Prawitt, D. F. (2014). Trust and financial reporting quality. *Journal of Accounting Research, 52*(5), 1087-1125.

Ghauri, P., & Grønhaug, K. (2010). *Research methods in business studies* (4th ed.). Financial Times Prentice Hall.

Giokas, D. I. (2008). Assessing the efficiency in operations of a large Greek bank branch network adopting different economic behaviors. *Economic Modelling, 25*(3), 559-574.

Giokas, D. I., Eriotis, N., & Dokas, I. (2015). Efficiency and productivity of the food and beverage listed firms in the pre-recession and recessionary periods in Greece. *Applied Economics, 47*(19), 1-15.

Giorgio, L. D., Flaxman, A. D., Moses, M. W., Fullman, N., Hanlon, M., Conner, R. O., Wollum, A., & Murray, C. J. L. (2016). Efficiency of health care production in low-resource settings: A Monte-Carlo simulation to compare the performance of Data Envelopment Analysis, stochastic distance functions, and an ensemble model. *PLoS ONE, 11*(1)

Girod, S. J. G., & Whittington, R. (2017). Reconfiguration, restructuring and firm performance: Dynamic capabilities and environmental dynamism. *Strategic Management Journal, 38*(5), 1121-1133.

Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega, 17*(3), 237-250.

Goodwin, C. J., & Goodwin, K. A. (2017). *Research in Psychology Methods and Design* (8th ed.). John Wiley & Sons.

Gul, F. A., Khedmati, M., Lim, E. K., & Navissi, F. (2018). Managerial ability, financial distress, and audit fees. *Accounting Horizons, 32*(1), 29-51.

Guo, J., Huang, P., Zhang, Y., & Zhou, N. (2016). The effect of employee treatment policies on internal control weaknesses and financial restatements. *The Accounting Review, 91*(4), 1167-1194.

Guzmán, I., Arcas, N., Ghelfi, R., & Rivaroli, S. (2009). Technical efficiency in the fresh fruit and vegetable sector: a comparison study of Italian and Spanish firms. *Fruits, 64*(4), 243-252.

Guzmán, I., & Arcas, N. (2008). The usefulness of accounting information in the measurement of technical efficiency in agricultural cooperatives. *Annals of Public and Cooperative Economics, 79*(1), 107-131.

Hadad, M. D., Hall, M. J. B., Kenjegalieva, K. A., Santoso, W. h., & Simper, R. (2011). Productivity changes and risk management in Indonesian banking: A Malmquist analysis. *Applied Financial Economics, 21*(12), 847-861.

Hamann, P. M., & Schiemann, F. (2021). Organizational performance as a set of four dimensions: An empirical analysis. *Journal of Business Research, 127*, 45-65.

Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). University of Chicago Press.

Harper, J., Kemper, K. J., & Sun, L. (2019). Managerial ability and bond rating changes. *Financial Markets, Institutions and Instruments, 28*(5), 381-401.

Harper, J., & Sun, L. (2019). Geographic distance and goodwill impairment. *International Journal of Accounting and Information Management, 27*(4), 547-572.

Harrison, J., & Rouse, P. (2016). DEA and accounting performance measurement. In S. N. Hwang, H. S. Lee & J. Zhu (Eds.), *Handbook of Operations Analytics Using Data Envelopment Analysis* (pp. 385-412). Springer.

Harrison, J., Rouse, P., & Armstrong, J. (2012). Categorical and continuous non-discretionary variables in Data Envelopment Analysis: a comparison of two single-stage models. *Journal of Productivity Analysis, 37*(3), 261-276.

Harrison, J., Rouse, P., & De Villiers, C. (2012). Accountability and performance measurement: A stakeholder perspective. *Journal of Centrum Cathedra, 5*(2), 243-258.

Hasan, M. M. (2020). Readability of narrative disclosures in 10-K reports: does managerial ability matter? *European Accounting Review, 29*(1), 147-168.

Hatami-Marbini, A., & Toloo, M. (2017). An extended multiple criteria Data Envelopment Analysis model. *Expert Systems with Applications, 73*, 201-219.

Hetico, H. R., & Marcinko, D. E. (2007). *Dictionary of Health Economics and Finance*. Springer.

Hillman, A. J., & Keim, G. D. (2001). Shareholder value, stakeholder management, and social issues: What's the bottom line? *Strategic Management Journal, 22*(2), 125-139.

Ho, C. B., Wu, D. D., & Olson, D. L. (2009). A risk scoring model and application to measuring internet stock performance. *International Journal of Information Technology and Decision Making, 8*(01), 133-149.

Hopper, T., & Powell, A. (1985). Making sense of research into organisational and social aspects of management accounting a review of its underlying assumptions. *Journal of Management Studies, 22*(5), 429-465.

Hopwood, A. G. (1987). The archaeology of accounting systems. *Accounting, Organizations and Society, 12*(3), 207-234.

Hsiao, H., Chang, H., Cianci, A. M., & Huang, L. (2010). First financial restructuring and operating efficiency: Evidence from Taiwanese commercial banks. *Journal of Banking and Finance, 34*(7), 1461-1471.

Hsiao, H., & Lin, M. (2013). Taiwan second financial restructuring and commercial bank productivity growth: M&As impact. *Review of Accounting and Finance, 12*(4), 327-350.

Hsiao, S. (2014). PTE, innovation capital and firm value interactions in the biotech medical industry. *Journal of Business Research, 67*(12), 2636-2644.

Hsu, H., & Khan, S. (2019). Chief accounting officers and audit efficiency. *Asian Review of Accounting, 27*(4),614-638.

Huang, L., Hsiao, H., Cheng, M., & Chang, S. (2008). Effects of financial reform on productivity change. *Industrial Management and Data Systems, 108*(7).

International Accounting Standards Board. (2016) *International Financial Reporting Standard 16 Leases.*

International Accounting Standards Board. (2018). *International Accounting Standards Board Conceptual Framework.*

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

International Accounting Standards Board. (2020). *International Accounting Standard 38 Intangible Assets.*

James, W. G. (1992). *Our Precious Metal: African Labour in South Africa's Gold Industry, 1970-1990*. Indiana University Press 1992.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*(2), 199-218.

Joo, S. J., Min, H., Kwon, I., & Kwon, H. (2010). Comparative efficiencies of specialty coffee retailers from the perspectives of socially responsible global sourcing. *The International Journal of Logistics Management, 21*(3), 490-509.

Joo, S. J., Nixon, D., & Stoeberl, P. A. (2011). Benchmarking with Data Envelopment Analysis: A return on asset perspective. *Benchmarking: An International Journal, 18*(4), 529-542.

Joo, S. J., Stoeberl, P. A., & Fitzer, K. (2009). Measuring and benchmarking the performance of coffee stores for retail operations. *Benchmarking: An International Journal, 16*(6), 741-753.

Joo, S. J., Stoeberl, P. A., & Kwon, I. G. (2007). Benchmarking efficiencies and strategies for resale operations of a charity organization. *Benchmarking: An International Journal, 14*(4), 455-464.

Joshi, R. N., & Singh, S. P. (2010). Estimation of total factor productivity in the Indian garment industry. *Journal of Fashion Marketing and Management: An International Journal, 14*(1), 145-160.

Jradi, S., Parmeter, C. F., & Ruggiero, J. (2021). Quantile estimation of stochastic frontiers with the normal-exponential specification. *European Journal of Operational Research,*

Jradi, S., & Ruggiero, J. (2019). Stochastic Data Envelopment Analysis: A quantile regression approach to estimate the production frontier. *European Journal of Operational Research, 278*(2), 385-393.

Jung, B., Lee, W., & Weber, D. P. (2014). Financial reporting quality and labor investment efficiency. *Contemporary Accounting Research, 31*(4), 1047-1076.

Kaffash, S., Matin, R. K., & Tajik, M. (2018). A directional semi-oriented radial DEA measure: An application on financial stability and the efficiency of banks. *Annals of Operations Research, 264*(1-2), 213-234.

Kao, C., & Hwang, S. (2008). Efficiency decomposition in two-stage Data Envelopment Analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research, 185*(1), 418-429.

Kao, C., & Liu, S. (2004). Predicting bank performance with financial forecasts: A case of Taiwan commercial banks. *Journal of Banking and Finance, 28*(10), 2353-2368.

Kapelko, M. (2017a). Dynamic versus static inefficiency assessment of the Polish meat-processing industry in the aftermath of the European Union integration and financial crisis. *Agribusiness, 33*(4), 505-521.

Kapelko, M. (2017b). Measuring productivity change accounting for adjustment costs: evidence from the food industry in the European Union. *Annals of Operations Research, 1*(278), 1-20.

Kapelko, M. (2018). Measuring inefficiency for specific inputs using Data Envelopment Analysis: Evidence from construction industry in Spain and Portugal. *Central European Journal of Operations Research, 26*(1), 43-66.

Kapelko, M., Alfons, O. L., & Spiro, E. S. (2015). Effect of food regulation on the Spanish food processing industry: A dynamic productivity analysis. *PLoS ONE, 10*(6), 1-16.

Kapelko, M., Horta, I. M., Camanho, A. S., & Lansink, A. O. (2015). Measurement of input-specific productivity growth with an application to the construction industry in Spain and Portugal. *International Journal of Production Economics, 166*, 64-71.

Kapelko, M., & Lansink, A. O. (2015). Technical efficiency and its determinants in the Spanish construction sector pre- and post-financial crisis. *International Journal of Strategic Property Management, 19*(1), 96-109.

Kapelko, M., & Lansink, A. O. (2017). Dynamic multi-directional inefficiency analysis of European dairy manufacturing firms. *European Journal of Operational Research, 257*(1), 338-344.

Kapelko, M., & Lansink, A. O. (2018). Managerial and program inefficiency for European meat manufacturing firms: A dynamic multidirectional inefficiency analysis approach. *Journal of Productivity Analysis, 49*(1), 25-36.

Kapelko, M., Lansink, A. O., & Stefanou, S. (2017a). Assessing the impact of changing economic environment on productivity growth: The case of the Spanish dairy processing industry. *Journal of Food Products Marketing, 23*(4), 384-397.

Kapelko, M., Lansink, A. O., & Stefanou, S. E. (2014). Assessing dynamic inefficiency of the Spanish construction sector pre- and post-financial crisis. *European Journal of Operational Research, 237*(1), 349-357.

Kapelko, M., Lansink, A. O., & Stefanou, S. E. (2016). Investment age and dynamic productivity growth in the Spanish food processing industry. *American Journal of Agricultural Economics, 98*(3), 946-961.

Kapelko, M., Lansink, A. O., & Stefanou, S. E. (2017b). Input-specific dynamic productivity change: Measurement and application to European dairy manufacturing firms. *Journal of Agricultural Economics, 68*(2), 579-599.

Kapelko, M., Lansink, A. O., & Stefanou, S. E. (2015). Analyzing the impact of investment spikes on dynamic productivity growth. *Omega, 54*, 116-124.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard - measures that drive performance. *Harvard Business Review, 70*(1), 71-79.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika, 30*(1), 81-93.

Khezrimotlagh, D., Cook, W. D., & Zhu, J. (2019). Number of performance measures versus number of decision-making units in DEA. *Annals of Operations Research, 1*(1), 1-34.

Kim, J. (1978). In Mueller C. W. (Ed.), *Factor Analysis: Statistical Methods and Practical Issues*. Sage Publications.

Ko, B., Kim, J., Choi, Y., Kil, K., & Lee, G. (2020). Enhancing the competitiveness of Korea's container shipping industry through structural improvements. *Maritime Policy and Management, 47*(1), 57-72.

Koester, A., Shevlin, T., & Wangerin, D. (2017). The role of managerial ability in corporate tax avoidance. *Management Science, 63*(10), 3285-3310.

Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities. In T. C. Koopmans (Ed.), *Activities Analysis of Production and Allocation*. Wiley.

Kozmetsky, G., & Yue, P. (1998). Comparative performance of global semiconductor companies. *Omega, 26*(2), 153.

Krippendorff, K. (2013). *Content Analysis: An Introduction to Its Methodology* (3rd ed.). SAGE.

Krishnan, G., & Wang, C. (2015). The relation between managerial ability and audit fees and going concern opinions. *Auditing: A Journal of Practice AND Theory, 34*(3), 139-160.

Kuang, Y. F., Liu, X. K., Paruchuri, S., & Qin, B. (2020). CFO social ties to non-CEO senior managers and financial restatements. *Accounting and Business Research*, 1-35.

Kubick, T. R., & Lockhart, G. B. (2016). Do external labor market incentives motivate CEOs to adopt more aggressive corporate tax reporting preferences? *Journal of Corporate Finance, 36*, 255-277.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions* (1st ed.). University of Chicago Press.

Kuo, K. C., Lu, W. M., & Dinh, T. N. (2020). Firm performance and ownership structure: Dynamic network Data Envelopment Analysis approach. *Managerial and Decision Economics, 41*(4), 608-623.

Kuo, K., Kweh, Q. L., Ting, I. W. K., & Azizan, N. A. (2017). Dynamic network performance evaluation of general insurance companies: An insight into risk management committee structure. *Total Quality Management and Business Excellence, 28*(5-6), 542-558.

Kweh, Q. L., Kuo, K., Wang, W., & Liu, H. (2015). Board independence, family control, and performance in Taiwanese listed semiconductor companies. *Hitotsubashi Journal of Economics*, 93-115.

Kweh, Q. L., Lu, W., Azizan, N. A., & Wang, W. (2014). The S in corporate social responsibility and performance of the US telecommunications industry. *Knowledge and Process Management, 21*(3), 151-158.

Kweh, Q. L., Lu, W., Nourani, M., & Ghazali@ Mohd Zain, M. H. (2018). Risk management and dynamic network performance: an illustration using a dual banking system. *Applied Economics, 50*(30), 3285-3299.

Kweh, Q. L., Lu, W., Ting, I. K., & Le, H. T. M. (2021). The cubic S-curve relationship between board independence and intellectual capital efficiency: Does firm size matter? *Journal of Intellectual Capital*, 1-27.

Kweh, Q. L., Lu, W., & Wang, W. (2014). Dynamic efficiency: Intellectual capital in the Chinese non-life insurance firms. *Journal of Knowledge Management, 18*(5), 937-951.

Kwon, H. B., Stoeberl, P. A., & Joo, S. J. (2008). Measuring comparative efficiencies and merger impacts of wireless communication companies. *Benchmarking: An International Journal, 15*(3), 212-224.

Lakshmanan, V. I., & Gorain, B. (2019). *Innovations and Breakthroughs in the Gold and Silver Industries Concepts, Applications and Future Trends*. Springer.

Laughlin, R. (1995). Methodological themes: Empirical research in accounting; alternative approaches and a case for 'middle-range' thinking. *Accounting, Auditing and Accountability, 8*(1), 63.

Law, K. S., & Wong, C. (1999). Multidimensional Constructs M Structural Equation Analysis: An Illustration Using the Job Perception and Job Satisfaction Constructs. *Journal of Management, 25*(2), 143-160.

Law, K. S., Wong, C., & Mobley, W. H. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Review, 23*(4), 741-755.

Lee, C. B., Wan, J., Shi, W., & Li, K. (2014). A cross-country study of competitiveness of the shipping industry. *Transport Policy, 35*, 366-376.

Lee, J. (2005). Using DEA to measure efficiency in forest and paper companies. *Forest Products Journal, 55*(1), 58-66.

Leightner, J. E., & Lovell, C. A. K. (1998). The impact of financial liberalization on the performance of Thai banks. *Journal of Economics and Business, 50*(2), 115-131.

Liu, H., Tsai, H., & Wu, J. (2018). Regional hotel performance and benchmarking in the Pearl River Delta. *International Journal of Contemporary Hospitality Management, 30*(2), 855-873.

Liu, J. S., Lu, L. Y. Y., Lu, W., & Lin, B. J. Y. (2012). A survey of DEA applications. *Omega, 41*(5), 893-902.

Lu, W., Wang, W., & Kweh, Q. L. (2014). Intellectual capital and performance in the Chinese life insurance industry. *Omega, 42*(1), 65-74.

Luo, X., & Donthu, N. (2005). Assessing advertising media spending inefficiencies in generating sales. *Journal of Business Research, 58*(1), 28-36.

Maccallum, R. C., & Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin, 114*(3), 533-541.

Mackenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *Journal of Applied Psychology, 90*(4), 710-730.

Magdaleno, M. I. A., & García, J.G. (2009). Cooperatives versus corporates in the Spanish agricultural sector: Non-parametric estimation of technical efficiency. *Anales De Estudios Económicos Y Empresariales,* (19), 61-90.

Mahajan, V., Nauriyal, D. K., & Singh, S. P. (2014). Technical efficiency analysis of the Indian drug and pharmaceutical industry: a non-parametric approach. *Benchmarking: An International Journal, 21*(5), 734-755.

Mahajan, V., Nauriyal, D. K., & Singh, S. P. (2015). Efficiency trends in the Indian pharmaceutical industry in the new patent regime. *International Journal of Business Performance Management, 16*(4), 389-406.

Mahajan, V., Nauriyal, D. K., & Singh, S. P. (2018). Efficiency and its determinants: Panel Data Evidence from the Indian Pharmaceutical Industry. *Margin: The Journal of Applied Economic Research, 12*(1), 19-40.

Majumdar, S. (1999). Comparative Organizational Characteristics of Indian State-Owned Enterprises. *Review of Industrial Organization; an International Journal Published for the Industrial Organization Society, 15*(2), 165-182.

Manandhar, R., & Tang, J. C. S. (2002). The evaluation of bank branch performance using Data Envelopment Analysis: A framework. *Journal of High Technology Management Research, 13*(1), 1-17.

Martín-Gamboa, M., Iribarren, D., García-Gusano, D., & Dufour, J. (2017). A review of life-cycle approaches coupled with Data Envelopment Analysis within multi-criteria decision analysis for sustainability assessment of energy systems. *Journal of Cleaner Production, 150*, 164-174.

Mazumdar, M. (2013). *Performance of Pharmaceutical Companies in India.* Springer Science and Business Media.

Mazumdar, M., & Rajeev, M. (2009). Comparing the efficiency and productivity of the Indian pharmaceutical firms: A Malmquist-meta-frontier approach. *International Journal of Business and Economics, 8*(2), 159-181.

McCallig, J., Robb, A., & Rohde, F. (2019). Establishing the representational faithfulness of financial accounting information using multiparty security, network analysis and a blockchain. *International Journal of Accounting Information Systems, 33*, 47-58.

McEachern, D., & Paradi, J. (2007). Intra- and inter-country bank branch assessment using DEA. *Journal of Productivity Analysis, 27*(2), 123-136.

McFadden, D. (Ed.). (1978). Cost, revenue, and profit functions, In Fuss, M. and McFadden, D. (Eds.), *Production Economics: A Dual Approach to Theory and Applications,* Volume 1, Amsterdam: North-Holland.

Mckee, G., & Kagan, A. (2018). Managerial implications of off-balance sheet items in community banks. *Studies in Economics and Finance, 35*(1), 178-195.

Mhatre, N., Joo, S. J., & Lee, C. C. (2014). Benchmarking the performance of department stores within an income elasticity of demand perspective. *Benchmarking: An International Journal, 21*(2), 205-217.

Miller, C. C., Washburn, N. T., & Glick, W. H. (2013). The Myth of Firm Performance. *Organization Science, 24*(3), 948-964.

Miller, P., & O'Leary, T. (1987). Accounting and the construction of the governable person. *Accounting, Organizations and Society, 12*(3), 235-265.

Miller, S. M., & Noulas, A. G. (1996). The technical efficiency of large bank production. *Journal of Banking and Finance, 20*(3), 495-509.

Min, H., & Joo, S. J. (2006). Benchmarking the operational efficiency of third party logistics providers using Data Envelopment Analysis. *Supply Chain Management: An International Journal, 11*(3), 259-265.

Min, H., & Joo, S. J. (2009). Benchmarking third-party logistics providers using Data Envelopment Analysis: An update. *Benchmarking: An International Journal, 16*(5), 572-587.

Min, H., Min, H., & Joo, S. (2008). A Data Envelopment Analysis-based balanced scorecard for measuring the comparative efficiency of Korean luxury hotels. *International Journal of Quality and Reliability Management, 25*(4), 349-365.

Minj, J., Aparna, S. V., & Kumari, A. (2020). *Dairy Processing Advanced Research to Applications*. Springer.

Nikiforos, M., & Zezza, G. (2017). Stock-flow consistent macroeconomic models: A survey. *Journal of Economic Surveys, 31*(5), 1204-1239.

Noulas, A. G., Glaveli, N., & Kiriakopoulos, I. (2008). Investigating cost efficiency in the branch network of a Greek bank: An empirical study. *Managerial Finance, 34*(3), 160-171.

Nourani, M., Devadason, E. S., Kweh, Q. L., & Lu, W. (2017). Business excellence: The managerial and value-creation efficiencies of the insurance companies. *Total Quality Management and Business Excellence, 28*(7-8), 879-896.

Nourani, M., Ting, I. W. K., Lu, W., & Kweh, Q. L. (2019). Capital structure and dynamic performance: Evidence from Asian-5 banks. *The Singapore Economic Review, 64*(03), 495-516.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. (3rd ed.). New York: McGraw-Hill.

Oberholzer, M. (2010). A model to compare firms' efficiency in creating shareholders' value. *South African Journal of Accounting Research, 24*(1), 13-30.

Oberholzer, M. (2012). The relative importance of financial ratios in creating shareholders' wealth. *South African Journal of Economic and Management Sciences, 15*(4), 416-428.

Oberholzer, M. (2014a). The relationship between company size and CEO remuneration: A scaling perspective. *Corporate Ownership and Control, 11*(4), 412-423.

Oberholzer, M. (2013). A non-parametric comparison among firms' income statement-based and balance sheet-based performance. *The International Business and Economics Research Journal (Online), 12*(11), 1467-1478.

Oberholzer, M. (2014b). A model to estimate firms' accounting-based performance: A Data Envelopment Approach. *The International Business and Economics Research Journal, 13*(6), 1301.

Oberholzer, M., Mong, D., & van Romburgh, J. (2017). Towards a new model to benchmark firms' operating efficiency: A Data Envelopment Analysis approach. *South African Journal of Accounting Research, 31*(3), 223-239.

Oberholzer, M., & Theunissen, M. (2012). Benchmarking of Johannesburg stock exchange CEO compensation. *International Business and Economics Research Journal, 11*(9), 1061-1076.

Oberholzer, M., & Van der Westhuizen, G. (2009). Estimating technical and scale efficiency in banks and its relationship with economic value added: A South African study. *South African Journal of Accounting Research, 23*(1), 67-86.

Oberholzer, M., & Van der Westhuizen, G. (2010). The relationship between economic value added and Data Envelopment Analysis based on financial statements: Empirical evidence of South African listed banks. *Management Dynamics: Journal of the Southern African Institute for Management Scientists, 19*(3), 29-41.

Ochola, P. (2017). A two-stage performance improvement evaluation of the insurance industry in Kenya: An application of Data Envelopment Analysis and Tobit regression model. Working paper.

Oh, S., & Shin, J. (2015). The impact of mismeasurement in performance benchmarking: A Monte Carlo comparison of SFA and DEA with different multi-period budgeting strategies. *European Journal of Operational Research, 240*(2), 518-527.

Oral, M., & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. *European Journal of Operational Research, 46*(3), 282-294.

Oral, M., Kettani, O., & Yolalan, R. (1992). An empirical study on analysing the productivity of bank branches. *IIE Transactions, 24*(5), 166-176.

Paradi, J. C., Rouatt, S., & Zhu, H. (2011). Two-stage evaluation of bank branch efficiency using Data Envelopment Analysis. *Omega, 39*(1), 99-109.

Paradi, J. C., Sherman, H. D., & Tam, F. K. (2018). *Data Envelopment Analysis in the Financial Services Industry: A Guide for Practitioners and Analysts Working in Operations Research Using DEA.* Springer.

Parkan, C., & Wu, M. (1999). Measuring the performance of operations of Hong Kong's manufacturing industries. *European Journal of Operational Research, 118*(2), 235-258.

Patton, M. Q. (2002). In Patton M. Q. (Ed.), *Qualitative Research and Evaluation Methods* (3rd ed.). Sage Publications.

Perelman, S., & Santín, D. (2009). How to generate regularly behaved production data? A Monte Carlo experimentation on DEA scale efficiency measurement. *European Journal of Operational Research, 199*(1), 303-310.

Petter, S., Straub, D., & Rai, A. (2007). Specifying formative constructs in information systems research. *MIS Quarterly, 31*(4), 623-656.

Phan, D. H. B., Tran, V. T., Nguyen, D. T., & Le, A. (2020). The importance of managerial ability on crude oil price uncertainty-firm performance relationship. *Energy Economics, 88*, 104778.

Piesse, J., & Thirtle, C. (2000). A stochastic frontier approach to firm level efficiency, technological change, and productivity during the early transition in Hungary. *Journal of Comparative Economics, 28*(3), 473-501.

Pille, P., & Paradi, J. C. (2002). Financial performance analysis of Ontario (Canada) Credit Unions: An application of DEA in the regulatory environment. *European Journal of Operational Research, 139*(2), 339-350.

Ponta, L., Raberto, M., Teglio, A., & Cincotti, S. (2018). An agent-based stock-flow consistent model of the sustainable transition in the energy sector. *Ecological Economics, 145*, 274-300.

Portela, M. C. A. S. (2014). Value and quantity data in economic and technical efficiency measurement. *Economics Letters, 124*(1), 108-112.

Porter, M. E. (1985). Creating and sustaining superior performance. *Competitive Advantage, 167*, 167-206.

Porter, M. E. (2011). *Competitive Advantage of Nations: Creating and Sustaining Superior Performance*. Simon and Schuster.

Qiu, B., Trapkov, S., & Yakoub, F. (2014). Do target CEOs trade premiums for personal benefits? *Journal of Banking & Finance, 42*(1), 23-41.

Raheman, A., Afza, T., Qayyum, A., & Bodla, M. A. (2008). Estimating total factor productivity and its components: Evidence from major manufacturing industries of Pakistan. *Pakistan Development Review, 47*(4), 677-694.

Rahman, M., Lambkin, M., & Hussain, D. (2016). Value creation and appropriation following M&A: A Data Envelopment Analysis. *Journal of Business Research, 69*(12), 5628-5635.

Rahman, M., Rodríguez-Serrano, M. Á, & Lambkin, M. (2018). Brand management efficiency and firm value: An integrated resource based and signalling theory perspective. *Industrial Marketing Management, 72*, 112-126.

Rahman, M., Rodríguez-Serrano, M. Á, & Lambkin, M. (2019). Advertising efficiency and profitability: Evidence from the pharmaceutical industry. *Industrial Marketing Management*, 89, 619-629.

Raith, A., Rouse, P., & Seiford, L. M. (2019). Benchmarking using Data Envelopment Analysis: Application to stores of a post and banking business. *Multiple Criteria Decision Making and Aiding*, 1-39. Springer.

Ramanathan, K. V. (1985). A proposed framework for designing management control systems in not-for-profit organizations. *Financial Accountability and Management, 1*(1), 75-92.

Ravitch, S. M., & Riggan, M. (2012). *Reason and Rigor: How Conceptual Frameworks Guide Research*. Sage Publications.

Ray, S. C., Walden, J., & Chen, L. (2020). Economic measures of capacity utilization: A nonparametric cost function analysis. *European Journal of Operational Research,*

Richard, P. J., Devinney, T. M., Yip, G. S., & Johnson, G. (2009). Measuring organizational performance: Towards methodological best practice. *Journal of Management, 35*(3), 718-804.

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.

Rouse, P., Harrison, J., & Chen, L. (2010). Data Envelopment Analysis: A practical tool to measure performance. *Australian Accounting Review, 20*(2), 165-177.

Rouse, P., & Tripe, D. (2016). Allocative and technical efficiency of New Zealand Banks. *Meditari Accountancy Research, 24*(4), 574-587.

Ruggiero, J. (1998). A new approach for technical efficiency estimation in multiple output production, *European Journal of Operational Research, 111*(2), 369-380.

Ruggiero, J. (1999). Efficiency estimation and error decomposition in the stochastic frontier model: A Monte Carlo analysis. *European Journal of Operational Research, 115*(3), 555-563.

Ruggiero, J. (2004). Data Envelopment Analysis with stochastic data. *Journal of the Operational Research Society, 55*(9), 1008-1012.

Ruggiero, J. (2007). A comparison of DEA and the stochastic frontier model using panel data. *International Transactions in Operational Research, 14*(3), 259-266.

Ryan, B., Scapens, R. W., & Theobald, M. (2002). *research method and methodology in finance and accounting*, Cengage Learning.

Sallehu, M. (2017). Do restructurings improve post-restructuring productivity? *Journal of Accounting and Finance, 17*(1), 103-123.

Santín, D., & Sicilia, G. (2017). Dealing with endogeneity in Data Envelopment Analysis applications. *Expert Systems with Applications, 68*, 173-184.

Saranga, H. (2009). The Indian auto component industry – Estimation of operational efficiency and its determinants using DEA. *European Journal of Operational Research 196*(2), 707-718.

Saranga, H., & Phani, B. V. (2009). Determinants of operational efficiencies in the Indian pharmaceutical industry. *International Transactions in Operational Research, 16*(1), 109-130.

Sarwar, B., Kutan, A., Ming, X., & Husnain, M. (2019). How do talented managers view dividend policy? Further evidence from Chinese equity market. *International Journal of Emerging Markets, 15*(3), 559-586.

Sathye, M. (2001). X-efficiency in Australian banking: An empirical investigation. *Journal of Banking and Finance, 25*(3), 613-630.

Sathye, M. (2003). Efficiency of banks in a developing economy: The case of India. *European Journal of Operational Research, 148*(3), 662-671.

Sathye, S., & Sathye, M. (2017). Do ATMs increase technical efficiency of banks in a developing country? Evidence from Indian banks. *Australian Accounting Review, 27*(1), 101-111.

Seiford, L. M., & Zhu, J. (1999). Profitability and marketability of the top 55 US commercial banks. *Management Science, 45*(9), 1270-1288.

Seiford, L. M., & Zhu, J. (2002). Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research, 142*(1), 16-20.

Sevcovic, D., Halicka, M., & Brunovsky, P. (2001). DEA analysis for a large structured bank branch network. *Central European Journal of Operations Research, 9*(4), 329-342.

Shephard, R. W. (1953). *Cost and Production Functions*. Princeton University Press.

Shephard, R. W. (1970). *Theory of Cost and Production Functions*. Princeton University Press.

Shin, H., & Park, S. (2020). The internal control manager and operational efficiency: Evidence from Korea. *Managerial Auditing Journal, 35*(7), 979-1006.

Shiu, A. (2002). Efficiency of Chinese enterprises. *Journal of Productivity Analysis, 18*(3), 255-267.

Smith, P. (1990). Data Envelopment Analysis applied to financial statements. *Omega, 18*(2), 131-138.

Soboh, R., Lansink, A. O., & Van Dijk, G. (2012). Efficiency of cooperatives and investor owned firms revisited. *Journal of Agricultural Economics, 63*(1), 142-157.

Soliman, M. T. (2008). The use of DuPont analysis by market participants. *The Accounting Review, 83*(3), 823-853.

Spanos, Y. E., Zaralis, G., & Lioukas, S. (2004). Strategy and industry effects on profitability: Evidence from Greece. *Strategic Management Journal, 25*(2), 139-165.

Spiggle, S. (1994). Analysis and interpretation of qualitative data in consumer research. *Journal of Consumer Research, 21*(3), 491-503.

Stamp, P. (Ed.). (1993). *In search of reality*. Routledge.

Stewart, T. J. (1992). A critical survey on the status of multiple criteria decision-making theory and practice. *Omega, 20*(5), 569-586.

Stewart, T. J. (1996). Relationships between Data Envelopment Analysis and multicriteria decision analysis. *Journal of the Operational Research Society, 47*(5), 654-665.

Sun, L. (2017). Managerial ability and chemical releases. *Sustainability Accounting, Management and Policy Journal, 8*(3), 281-306.

Thanassoulis, E. (1997). Duality in Data Envelopment Analysis under constant returns to scale. *IMA Journal of Management Mathematics, 8*(3), 253-266.

Tone, K., Kweh, Q. L., Lu, W., & Ting, I. W. K. (2019). Modeling investments in the dynamic network performance of insurance companies. *Omega, 88*, 237-247.

Truong, C., Garg, M., Adrian, C., Pham, A. V., & Shane, P. (2020). Political alignment and audit pricing. *International Journal of Auditing, 24*(2), 205-231.

Tsolas, I. E. (2011). Modelling profitability and effectiveness of Greek-listed construction firms: An integrated DEA and ratio analysis. *Construction Management and Economics, 29*(8), 795-807.

Venkatraman, N., & Ramanujam, V. (1986). Measurement of business performance in strategy research: A comparison of approaches. *Academy of Management Review, 11*(4), 801-814.

Walters, B. A., Le, S. A., & Kroll, M. (2015). Post-IPO governance and top management team rent generation and appropriation. *Journal of Business Research, 68*(1), 47-55.

Wan, Z., Zhu, M., Chen, S., & Sperling, D. (2016). Pollution: Three steps to a green shipping industry. *Nature News, 530*(7590), 275.

Wang, C. (2020). Debt maturity choice in CEOs' incentive to signal abilities. *Asia-Pacific Journal of Accounting and Economics,*1-17.

Wang, J. (2006). Corporate performance efficiency investigated by Data Envelopment Analysis and balanced scorecard. *The Journal of American Academy of Business, 9*(2), 312-318.

Wang, W., Lin, F., Ting, I. W. K., Kweh, Q. L., Lu, W., & Chiu, T. (2017). Does asset-light strategy contribute to the dynamic efficiency of global airlines? *Journal of Air Transport Management, 62*, 99-108.

Wang, W., Lu, W., Kweh, Q. L., & Lai, H. (2014). Does corporate social responsibility influence the corporate performance of the US telecommunications industry? *Telecommunications Policy, 38*(7), 580-591.

Wang, W., Lu, W., Kweh, Q. L., & Liu, Y. (2017). Decentralized and concentrated investments in China and the performance of Taiwanese listed electronic companies. *Applied Economics, 49*(25), 2443-2455.

Wang, W., Lu, W., Ting, I. W. K., & Kao, H. (2020). Asset-light strategy, managerial ability, and corporate performance of the Asian telecommunications industry. *Managerial and Decision Economics, 41*(8), 1568-1579.

Wang, W., Ting, I. W. K., Kuo, K., Kweh, Q. L., & Lin, Y. (2018). Corporate diversification and efficiency: evidence from Taiwanese top 100 manufacturing firms. *Operational Research, 18*(1), 187-203.

Wang, Z., Chen, M., Chin, C. L., & Zheng, Q. (2017). Managerial ability, political connections, and fraudulent financial reporting in China. *Journal of Accounting and Public Policy, 36*(2), 141-162.

Wanke, P., Tsionas, M. G., Chen, Z., & Antunes, J. J. M. (2020). Dynamic network DEA and SFA models for accounting and financial indicators with an analysis of super-efficiency in stochastic frontiers: An efficiency comparison in OECD banking. *International Review of Economics & Finance, 69*, 456-468.

Warner, W. L., Meeker, M., & Eells, K. (1949). *Social Class in America; a Manual of Procedure for the Measurement of Social Status.* Science Research Associates.

Wen, W., Cui, H., & Ke, Y. (2020). Directors with foreign experience and corporate tax avoidance. *Journal of Corporate Finance, 62*, 101624.

Weygandt, J. J., Kimmel, P. D., & Kieso, D. E. (2015). *Financial Accounting: IFRS edition* (3rd ed.). Wiley.

Williams, P. F. (2014). The myth of rigorous accounting research. *Accounting Horizons, 28*(4), 869-887.

Worthington, A. C. (2004). Determinants of merger and acquisition activity in Australian cooperative deposit-taking institutions. *Journal of Business Research, 57*(1), 47-57.

Worthington, A. C. (1998). The application of mathematical programming techniques to financial statement analysis: Australian gold production and exploration. *Australian Journal of Management, 23*(1), 97.

Worthington, A. C., & Hurley, E. V. (2002). Cost efficiency in Australian general insurers: A non-parametric approach. *The British Accounting Review, 34*(2), 89-108.

Wu, Y., Ting, W. K. I., Lu, W., Nourani, M., & Kweh, Q. L. (2016). The impact of earnings management on the performance of ASEAN banks. *Economic Modelling, 53*, 156-165.

Yang, T. H., & Choi, H. P. (2013). Performance analysis of international steel manufactures a benchmark study for steel supply chains. *International Journal of Supply Chain Management, 2*(3), 25-31.

Yang, Y. J., Chen, J., Kweh, Q. L., & Hsin, C. C. (2013). Ownership structure and efficiency in Taiwanese electronics firms. *Review of Accounting and Finance, 12*(4), 351-368.

Yeh, Q. (1996). The application of Data Envelopment Analysis in conjunction with financial ratios for bank performance evaluation. *Journal of the Operational Research Society, 47*(8), 980.

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. *Industrial Marketing Management, 43*(1), 25-31.

Yusof, K. N. C. K., Razali, A. R., & Tahir, I. M. (2010). An evaluation of company operation performance using Data Envelopment Analysis (DEA) approach: A study on Malaysian public listed companies. *International Business Management, 4*(2), 47-52.

Zelenyuk, V. (2020). Aggregation of inputs and outputs prior to Data Envelopment Analysis under big data. *European Journal of Operational Research, 282*(1), 172-187.

Zelenyuk, V., & Zheka, V. (2006). Corporate governance and firm's efficiency: The case of a transitional country, Ukraine. *Journal of Productivity Analysis, 25*(1), 143-157.

Zhang, W. D., Zhang, S., & Luo, X. (2006). Technological progress, inefficiency, and productivity growth in the US securities industry, 1980–2000. *Journal of Business Research, 59*(5), 589-594.

Zhou, G., Min, H., Xu, C., & Cao, Z. (2008). Evaluating the comparative efficiency of Chinese third-party logistics providers using Data Envelopment Analysis. *International Journal of Physical Distribution and Logistics Management, 38*(4), 262-279.

Zhu, J. (2000). Multi-factor performance measure model with an application to Fortune 500 companies. *European Journal of Operational Research, 123*(1), 105-124.