New Insights on Performance Assessment and Improvement in the Construction Industry

Isabel Maria Noronha de Resende Horta e Costa

A Thesis submitted to Faculdade de Engenharia da Universidade do Porto
for the doctoral degree in Industrial Engineering and Management

Supervisors

Professor Ana Maria Cunha Ribeiro dos Santos Ponces Camanho
Professor Jorge Manuel Fachana Moreira da Costa
Professor Rui Manuel Soucasaux Meneses e Sousa

Faculdade de Engenharia da Universidade do Porto
2011

Abstract

In recent years, the competitive environment of the Construction Industry (CI) has become increasingly fierce. The companies are aware of the challenges imposed, and have attempted to implement systematic methods of performance assessment in order to achieve competitive advantage. This thesis intends to develop a robust approach for performance assessment and improvement in construction companies based on Data Envelopment Analysis (DEA). The research includes four main research topics described next.

The first topic develops a model to assess the performance of construction companies based on a set of evaluation criteria. This model was designed to be integrated in e-marketplaces to comply with the recent technological advances in the CI. The model functions as a decision support system to aid the selection of the company to be contracted among competitive bids.

The second topic proposes a methodology for assessing company performance using DEA to complement the information provided by key performance indicators (KPIs) available in web benchmarking platforms, which are tools frequently used by construction companies. This methodology provides managerial insights concerning overall performance and improvement targets. Enhanced DEA models were proposed to enable a more realistic assessment of companies and to suggest targets to all organizations, even for the best-practice companies, informed by decision maker preferences.

The third topic explores trends in the performance of the Portuguese CI, and identifies the factors that promote excellence and innovation in the sector. It is also proposed an enhanced model to assess innovation within an industry, identifying the innovative companies and quantifying the extent of innovation. This research enhances the construction of composite indicators using DEA. The determinants of good performance and innovation are explored using regression techniques. The statistical significance of the results is ensured by the use of bootstrapping.

The last topic investigates the efficiency level of construction companies worldwide, exploring in particular the effect of location and activity in efficiency. In addition, valuable insights are provided concerning the convergence in efficiency across regions. DEA is used to estimate efficiency, and the Malmquist index is applied for the evaluation of productivity change over time. Both methods were complemented by bootstrapping to refine the estimates obtained. A panel data truncated regression was used to explore the impact of location and activity in the efficiency levels.

In summary, this thesis aims to show the usefulness of DEA combined with other techniques to guide companies in the definition of optimal strategies for continuous performance improvement.

Resumo

A competitividade da Indústria da Construção (IC) tem vindo a intensificar-se em todo o mundo. Estes aspetos têm reforçado a necessidade das empresas da construção adotarem metodologias de avaliação sistemática do desempenho visando alcançar vantagem competitiva. Esta tese pretende desenvolver uma metodologia robusta de avaliação e melhoria do desempenho das empresas da IC, baseada na técnica de *Data Envelopment Analysis* (DEA). A presente investigação divide-se em quatro partes principais:

A primeira parte desenvolve um sistema de avaliação do desempenho para as empresas da IC incluindo um conjunto alargado de indicadores de desempenho. Este sistema foi desenvolvido para ser integrado em plataformas de comércio eletrónico, dada a crescente expansão destas plataformas no sector. O sistema de apoio à decisão visa facilitar a seleção da empresa a contratar.

A segunda parte propõe uma metodologia de avaliação do desempenho utilizando a técnica DEA para complementar a informação fornecida por indicadores de desempenho disponíveis através de plataformas de *benchmarking*, que são frequentemente utilizadas pelas empresas da IC. Esta metodologia fornece informação relevante relacionada com o desempenho global da empresa e com a definição de metas a alcançar. Os modelos de DEA foram adaptados para avaliarem as empresas de uma forma mais realista e fornecerem metas para todas as empresas, mesmo para as eficientes.

A terceira parte explora os níveis de desempenho das empresas da IC em Portugal e identifica fatores que promovem uma melhoria do desempenho. É ainda proposto um modelo para identificar empresas inovadoras e quantificar o nível de inovação. Esta investigação tem por base indicadores compósitos determinados através da técnica de DEA para avaliar o desempenho e técnicas de regressão para identificar os fatores que promovem excelência e inovação. A relevância estatística dos resultados é assegurada pelo uso de *bootstrapping*.

A última parte investiga os níveis de eficiência das empresas da IC à escala mundial e explora, em detalhe, o efeito da localização e da atividade da empresa na eficiência. É ainda proposta uma metodologia que permite avaliar a convergência do nível de eficiência em diferentes regiões. O DEA é utilizado para determinar a eficiência das empresas e o índice de *Malmquist* para explorar as mudanças de desempenho ao longo do tempo. O impacto da localização e da atividade na eficiência é avaliado através de uma regressão truncada para dados em painel.

Em síntese, esta tese pretende demonstrar a utilidade do DEA combinado com outras técnicas no sentido de ajudar as empresas a delinear estratégias sustentadas de melhoria contínua do seu desempenho.

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Ana Camanho, for her excellent guidance, knowledge transmitted, constructive criticism, and constant encouragement, which were invaluable for the completion of this thesis. It was a privilege to work with such a professional and inspiring person.

I wish to thank Professor Jorge Moreira da Costa, for opening the doors for a fruitful research carrier, and for his example of perseverance throughout the time we worked together.

I would also want to thank the Professors from the Department of Industrial Engineering and Management, with whom I worked closer for always encouraging me to pursue my doctoral research. I am specially grateful to Professor Henriqueta Nóvoa for her continuous interest and encouragement.

I am particularly grateful to my parents, my brothers, and my nephew for all the motivation and the driving force throughout the time of this research. Also, I would like to thank my friends, specially my great friend Inês. Most of all, I wish to thank my mother Isabel for being an admirable example of life, her happiness, dedication and love made this work possible.

I would like to express my gratitude to my husband Alexandre, to whom I dedicate this work, for all the understanding, never-ending support, and ever-present love.

The funding of this research provided by Fundação para a Ciência e a Tecnologia (FCT) is also gratefully acknowledged.

Glossary

CI: Construction Industry

CRS: Constant Returns to Scale

DEA: Data Envelopment Analysis

DMU: Decision Making Unit

EC: Efficiency Change

GC: General Contractor

GDP: Gross Domestic Product

KPI: Key Performance Indicator

MI: Malmquist Index

PPS: Production Possibility Set

SC: Subcontractor

TC: Technological Change

VRS: Variable Returns to Scale

Table of Contents

List of Figures

Chapter 2

Chapter 4

Chapter 5

List of Tables

Chapter 6

Chapter 7

Chapter 1

Introduction

1.1 General context

Construction is a major industry worldwide accounting for a sizeable proportion of most countries gross domestic product (GDP). The global construction industry (CI) makes up approximately 9% of the world's GDP. The sector is the largest industrial employer in most countries, accounting to around 7% of the total employment worldwide, and almost half of the total resources used over the world (CSIR, 2006). The importance of the industry is related not only to its size but also to its outcome, which supports many other economic activities and contributes to socio-economic development of societies.

In recent years, the construction industry has been witnessing major structural changes, such as globalization, technological evolution, increased regulation, and a growing importance in the economy of both developed and developing countries. This contributed to a considerable increase in competition among construction companies.

Globalization has become an unavoidable fact in today's construction activity due to the recent developments in transport and communication, coupled with the creation of protocols, promoted by the World Trade Organization, enabling access to markets previously isolated. One of the most important changes in the CI attributable to globalization is the emergence of business opportunities for contractors to expand into new foreign markets. Construction companies, mainly from the developed countries, are adopting strategies of internationalization that enable them to benefit from the global market. For instance, some American and European companies are expanding their operations to Asian countries, with lower running costs and ample business opportunities. In developing countries, construction companies are capturing technological, financial and managerial know-how from international companies, narrowing the gap between both. In this truly global market, construction companies should be prepared to compete at both national and international level.

The construction industry is a very fragmented industry with a huge proportion of small companies, and is driven by unique construction projects with specific teams made up of varying combinations of different companies. The construction projects are then typically characterized by the involvement of many agents, including the owner, architectural and engineering companies, general contractors, subcontractors, construction materials' suppliers and producers. In addition, the construction industry is a labor intensive sector with low qualified labor force. No training is provided due to the cyclical activity of the CI that implies a significant rotation of workers. Due to the multi-relational, multi-dimensional and multi-disciplinary nature of the CI, it is difficult to implement in-depth technological advances. However, a few capable contractors are benefitting from technological advantages, and encouraging the spread of cutting edge technology in the CI. Some examples include modern design methods, such as the building information model and the multi-dimensional computer-aided design, or computer technologies, such as information systems and project web sites. The use of technological innovations is essential for pacing up the work in the sector.

The increased regulation of the construction industry activity worldwide has contributed to a significant change in the way of working and partnering in the CI. This lead to improvements in transparency in the industry, and also to increased competitiveness among construction companies. The bidding process for public sector projects is increasingly governed by local and national regulations, often subject to supranational rules, such as the Europe Union Public Procurement Directives, or World Trade Organization rules. This implies that contracts are increasingly being won based on evaluations considering several competitive factors, such as company technical capability or financial stability, complementing the price criterion. The level of regulation in the sector should be stringent as poor quality construction can be costly to owners and potentially hazardous to clients and the environment.

The challenging and highly competitive environment of the construction industry has caused performance improvement to be an increasingly important target. The construction companies are aware of the challenges imposed by this context and attempt to implement systematic methods to measure performance and search for best-practices to achieve competitive advantage and prosperity in the long-run. The topic of performance improvement is also of particular interest to government planners in order to encourage excellence in the sector, which is essential to foster economic development. This thesis addresses the issues of performance measurement and improvement in construction companies, providing insights both for managers and government planners. Due to the importance of the construction industry in the world economy, coupled with the substantial changes it has recently experiencing, this industry was chosen as the underlying context for the research of this thesis.

1.2 Motivation and research objectives

The evaluation of construction company performance typically involves the use of key performance indicators frequently available in web benchmarking platforms. The fierce competition in the construction industry motivated companies to strive for efficiency and competitiveness in order to survive and prosper in the market. Despite the considerable amount of research related to performance measurement in the construction industry in recent years, the performance measurement methods currently available are not enough to provide the managerial information required for performance improvement in today's business context. More advanced and robust methods for benchmarking and dissemination of best practices should be developed and made available to construction companies.

This research proposes the use of frontier analysis methods to benchmark the relative performance of companies. Frontier methods enable comparisons with the best observed performance instead of average industry values. The major strengths of such methods are to provide a single overall measure of performance that would not be available to companies otherwise, and to allow the identification of areas of improvement and targets. The ability to demonstrate and quantify issues that decision makers' might only know in a general and qualitative way makes these methods particularly valuable in the CI. From the alternative frontier methods available, Data Envelopment Analysis (DEA) was chosen for the assessment and improvement of performance in construction companies due to the greater flexibility to incorporate the multidimensional nature of the CI activity, captured by multiple inputs and outputs, and the use of minimal assumptions on the shape of the best-practice frontier.

The main purpose of this thesis is to develop a robust approach for performance assessment and improvement in construction companies. It involves the development of innovative models and methodologies applicable both at organizational and industry level. The research starts by developing models to evaluate performance at a company wide level and to provide insights concerning the strengths, weaknesses and targets for improvement (chapters 4 and 5). This is followed by the development of models to explore performance trends in the construction industry and to identify the factors that promote performance improvement (chapters 6 and 7).

The specific objectives of the research described in this thesis are as follows:

- To design a performance assessment system based on multiple criteria, adapted to e-marketplaces in the construction industry (chapter 4).
- To combine the use of Data Envelopment Analysis and Key Performance Indicators available in web benchmarking platforms for the assessment of performance at the company level (chapter 5).
- To assess financial soundness and innovation of Portuguese construction companies, identifying the factors that promote performance improvement (chapter 6).
- To measure the efficiency level and productivity change of construction companies worldwide, exploring the effect of location and activity in the performance levels (chapter 7).

Throughout the thesis, it was made an effort to ensure that the models and methodologies developed in the context of the construction industry can be easily applied to other organizations from different activity sectors. The use of innovative models for benchmarking best practices and improving performance is crucial to all organizations in order to strength the competitive position and to guarantee viability in today's challenging world.

1.3 Thesis summary

This thesis is divided in eight chapters which can be synthesized as follows. Chapter 2 presents an introduction to the DEA technique, giving particular attention to the most important extensions that are essential to pursue the objectives of this thesis.

Chapter 3 reviews the literature on performance measurement in the construction industry, with particular focus on the recent developments concerning the application of DEA in this sector. This chapter also identifies gaps in the literature and explains how they were addressed in the empirical part of this thesis.

Chapter 4 presents a performance assessment system for construction companies, named CIsea, based on several criteria. A set of KPIs was developed covering three different perspectives: the financial performance, the operation performance, and the bid attributes. To comply with the major technological advances in the CI, this system was designed to be integrated in e-marketplaces. CIsea aims to facilitate the selection of the best company to be contracted among competitive bids, and incorporates additional features, such as to allow bilateral evaluations between companies.

Chapter 5 develops a methodology for performance assessment applying DEA to complement the information provided by a set of KPIs. To enable a more realistic assessment of CI companies, two types of DEA models are used: one allows factor weights to vary freely and the other includes weight restrictions. These models estimate an efficiency score for each organization, identifying efficient organizations and providing performance improvements targets for the others. To enable suggesting targets for all organizations, expert opinion is used to specify virtual companies, which are included in the efficiency assessment to define a practical frontier located beyond the productivity levels of the original DEA frontier. Based on a sample of 20 Portuguese leading contractors, the Portuguese web benchmarking system for the CI, icBench, is used to demonstrate the advantages of integrating the DEA method with KPI benchmark scores.

Chapter 6 explores the trends in the performance of the Portuguese construction industry, and identifies the factors that promote excellence and innovation in the sector. From a methodological perspective, this chapter enhances the construction of composite indicators using the principals of “benefit of the doubt” weighting. This involves the use of DEA to estimate weights for aggregating the KPIs of the construction companies. The determinants of good performance and innovation are explored using regression techniques. In addition, the statistical significance of the results is ensured by the use of bootstrapping. This chapter also proposes an enhanced methodology to assess innovation within an industry, identifying the innovative companies and quantifying the extent of innovation.

Chapter 7 presents an exploratory study to assess the efficiency level of construction companies worldwide, exploring in particular detail the effect of location and activity on the efficiency levels. This chapter also provides insights concerning the convergence in efficiency across regions. The companies are divided in three regions (Europe, Asia and North America), and in the three main construction activities (Buildings, Heavy Civil and Specialty Trade). DEA is used to estimate efficiency, and the Malmquist index is applied for the evaluation of productivity change. Both methods were complemented with bootstrapping to refine the estimates obtained. A panel data truncated regression with categorical regressors is used to explore the impact of location and activity in the efficiency levels.

Chapter 8 oversees the main contributions of this thesis and proposes directions for future research.

Chapter 2

The assessment of performance using DEA

2.1 Introduction

This chapter aims to provide an overview on Data Envelopment Analysis (DEA) as it will be the technique used in this thesis for the evaluation of efficiency in organizations. First of all, a brief historical review on the measurement of efficiency will be presented.

An organization typically uses several inputs to produce several outputs. The inputs correspond to the resources used, whereas the outputs are the products or services obtained from the production process. The relationship between the level of outputs produced and the level of inputs used can be expressed by a production function, which defines the maximum output attainable given a set of inputs (corresponding to a theoretical frontier). However, the production function is usually unknown so it needs to be derived empirically from a set of homogeneous organizations under assessment. The efficiency of an organization is defined by comparing its inputs and outputs to those of the best performing from its peers.

The seminal work of Cobb and Douglas (1928) related to the estimation of an average production function, contributed considerably to the development of this field in economics. Since then, more flexible production functions were developed based on empirical data, both at macroeconomic and microeconomic level. Although the estimation of average production functions has become common practice in economics, the estimation of efficiency only attracted widespread attention more recently due to the difficulty in estimating the theoretical frontier.

The literature on efficiency measurement became more widely acknowledged with Debreu (1951), Koopmans (1951), and Farrell (1957). Koopmans was the first to define the concept of technical efficiency, and Debreu provided the first measure of efficiency, called the "coefficient of resource utilization". Farrell extended previous works by proposing to estimate an empirical frontier against which actual efficiency could be compared. In particular, Farrell suggested changing the focus from absolute to relative efficiency by promoting the comparison of a unit to the best actually achieved by peers performing a similar function. This was a major contribution to enhance the traditional economic approach to the estimation of production functions.

After the seminal work of Farrell (1957), efficiency measurement methods evolved, leading to two distinct research lines that differ in the way of estimating the frontier: Data Envelopment Analysis developed by Charnes et al. (1978), and Stochastic Frontier Analysis developed by Aigner et al. (1977).

Stochastic Frontier Analysis is a parametric technique, such that the production function is specified using a mathematical form, usually the Cobb-Douglas or translog function. These functional forms are specified *a-priori*, and their parameters are estimated from the empirical data. Stochastic Frontier Analysis assumes that deviations from the estimated frontier are composed by inefficiency and random error. These deviations are determined using statistical techniques.

DEA is a nonparametric technique, such that the empirical production function is defined by a set of assumptions that the points in the production possibility set (PPS) are assumed to satisfy (see section 2.2.2 for more details). The frontier is formed by piecewise linear segments that connect the set of frontier observations. These observations dominate all the others, i.e. no other observation achieves simultaneously a better score in the dimensions analyzed. This technique is deterministic (as opposed to a stochastic) assuming that all deviations of observed production from the estimated frontier are exclusively explained by inefficiency. The efficiency measure is calculated using mathematical programming techniques. The DEA technique is described in more detail in section 2.2.

One of the major strengths of DEA is the lack of parameterization, i.e. it requires no *a-priori* assumptions about the form of the empirical production function, allowing greater flexibility in the assumptions imposed to the frontier. The major drawback of the DEA technique is to assume that no random factors affect the construction of the frontier, such as random noise or measurement errors in the data. To overcome this limitation, a research line has been pursued to develop the statistical foundations for DEA. First, Banker (1993) provided a theoretical foundation for statistical hypothesis testing in DEA. More recently, Simar and Wilson (1998, 1999, 2000, 2007) proposed the use of bootstrapping for nonparametric envelopment estimators, such as the DEA efficiency measure. The idea underlying the bootstrapping procedure is to approximate the sampling distribution of interest by simulating, or mimicking, the data generation process. This enables the estimation of unbiased point estimates and the construction of confidence intervals. Bootstrapping procedures are presented in sections 2.2.5 and 2.3.2.

The remainder of this chapter is structured as follows. Section 2.2 provides a brief review of the DEA technique, and presents the most important extensions that are relevant to accomplish the objectives of this thesis. Section 2.3 introduces the Malmquist index for productivity measurement over time. Section 2.4 summarizes and concludes.

2.2 Data Envelopment Analysis

2.2.1 Introduction to efficiency assessment

DEA is a technique to measure the efficiency of a relatively homogenous set of organizational units, such as companies, schools or hospitals. An efficiency measure compares the ratio output over input of the unit assessed with the maximum value of this ratio observed in the other units. This notion of efficiency leads to an easy evaluation in the case of analysis involving a single input and a single output, since it reduces to a comparison of a ratio (output/input) for the unit analyzed ($unit_{j_o}$), with the maximum value of this ratio observed in other units ($j = 1, \ldots, n$).

$$\text{Efficiency} = \frac{output_{j_o}/input_{j_o}}{\max_j output_j/input_j} \tag{2.1}$$

However, more typically processes and organizational decision making units (DMUs) use multiple inputs (resources) to produce multiple outputs (outcomes), as shown schematically in Figure 2.1. The selection of the inputs and outputs is vital in real-life assessments. The inputs should reflect the resources that affect the outputs, and the outputs should capture the relevant outcomes on which we wish to evaluate the DMU.

To be able to compute an efficiency measure in these circumstances, expression (2.1) must be generalized. This requires an aggregation of the multiple inputs ($i = 1, \ldots, m$) and outputs ($r = 1, \ldots, s$) in a single efficiency ratio,

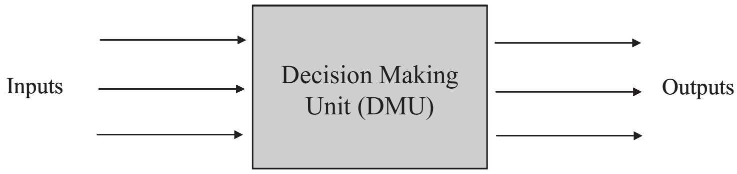

Figure 2.1: Schematic representation of a DMU used in a DEA assessment

corresponding to the weighted sum of outputs divided by the weighted sum of inputs. In this case, the computation of efficiency for a DMU_{j_o} under analysis would require estimating expression (2.2):

$$\text{Efficiency} = \frac{\sum_r u_r y_{rj_o} / \sum_i v_i x_{ij_o}}{\max_j \sum_r u_r y_{rj} / \sum_i v_i x_{ij}} \tag{2.2}$$

In expression (2.2), u_r and v_i are the weights attached to the outputs and inputs, respectively. The problem of this definition is that it requires a set of weights to be defined, and this can be difficult, particularly if a common set of weights to be applied across the set of organizational units is sought. A consensus regarding the selection of weights would certainly be difficult to achieve. As put forward by Charnes et al. (1978), this problem can be solved by arguing that individual units may have their own value systems, and therefore may legitimately define their own set of weights, such that their efficiency is maximized in the comparison with all other units in the sample. This is not only a sound of economic justification for the weights assigned, as the weights are allowed to reflect the value system and strategic options of each DMU, but also no inefficient unit can complain that its score would have been better if a different set of weights were used. The identification of the weights for each DMU_{j_o}, that show it in the best possible light, involves estimating the maximum value of the ratio defined in (2.3), where the weights are decision variables not known a priori (Cooper et al., 1996).

$$\text{Efficiency} = \underset{u_r,v_i}{max} \frac{\sum\limits_r u_r y_{rj_o} \Big/ \sum\limits_i v_i x_{ij_o}}{\max_j \sum\limits_r u_r y_{rj} \Big/ \sum\limits_i v_i x_{ij}} \tag{2.3}$$

Charnes et al. (1978) showed how to obtain the solution to this problem of finding the best possible weights for each unit, using mathematical programming techniques. The original mathematical programming model proposed for estimating relative efficiency, presented in the seminal paper by Charnes et al. (1978) is shown in (2.4):

$$\begin{aligned}
& e_{j_o} = \max \frac{\sum\limits_{r=1}^{s} u_r \ y_{rj_o}}{\sum\limits_{i=1}^{m} v_i \ x_{ij_o}} \\
& \text{subject to} \\
& \frac{\sum\limits_{r=1}^{s} u_r \ y_{rj}}{\sum\limits_{i=1}^{m} v_i \ x_{ij}} \leq 1, \ j = 1, \ldots, n \\
& v_i \geq \epsilon, \ i = 1, \ldots, m \\
& u_r \geq \epsilon, \ r = 1, \ldots, s
\end{aligned} \tag{2.4}$$

This model searches for the optimal input and output weights that maximize the efficiency of DMU_{j_o} under assessment, subject to the condition that the efficiency of all units in the sample is less than or equal to 1, when evaluated with the same set of weights. The other two constraints are included to guarantee that weights are positive and higher than a very small number ϵ, to take into account all the inputs and outputs in the efficiency assessment.

As a result, the efficiency measure ($e^*_{j_o}$) of DMU_{j_o}, obtained at the optimal solution to the DEA model, is between 0 and 1. The symbol ($*$) denotes the value of a variable at the optimal solution. The efficient DMUs obtain a performance score equal to 1, and the inefficient ones obtain a score lower than 1. The efficient DMUs are considered as examples of best practices (or benchmarks), and are used to specify the efficient frontier. For the inefficient DMUs, the magnitude of their inefficiency is derived by the distance to the frontier constructed from the benchmark DMUs. This comparison with benchmarks also allows determining the input and output targets corresponding to efficient operation.

The basic ideas behind DEA are explained using an illustrative example, using two outputs (y_1) and (y_2), and one input (x). Figure 2.2 illustrates the production possibility set, with the outputs normalized by the inputs (y_1/x) and (y_2/x) to allow a two-dimensional representation of this example with an output oriented perspective.

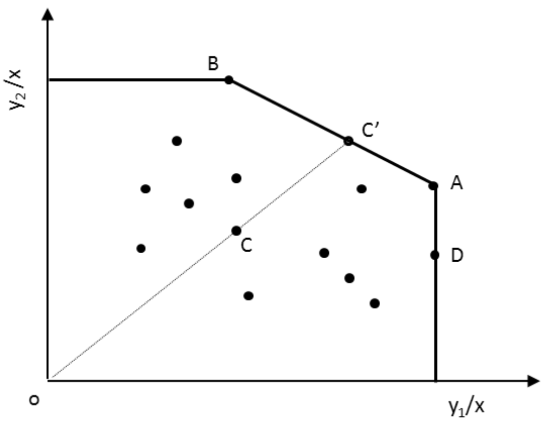

Figure 2.2: Representation of the production possibility set

From Figure 2.2, it is possible to visualize that the frontier of the production possibility set (PPS) is defined by the segment linking DMUs A and B, and the extensions parallel to the axes spanning from A and B. DMUs A and B define the efficient frontier that envelops all other DMUs. Therefore, they both achieve an efficiency score equal to 1 (or 100%) in the DEA models.

This means that these DMUs dominate all the others, i.e. no other DMU achieves simultaneously a better score in both dimensions. In particular, B achieves the highest value of the indicator (y_2/x), whereas A is the best on (y_1/x).

The information for inefficient units provided by a DEA analysis is illustrated using DMU C. The efficiency score of company C is graphically represented by the ratio OC/OC', which corresponds to a value lower than one. The output targets that would render C efficient correspond to point C'. DMUs A and B are the benchmarks for DMU C, and the targets for C can be obtained by a linear combination of these two DMUs.

Finally, it is important to distinguish between two concepts of efficiency proposed by Farrell (1957) and Koopmans (1951) as they differ for any DMU on an expansion of the frontier parallel to the axes. According to Farrell's efficiency notion, a DMU is technically efficient if it is not possible to increase the outputs (or decrease the inputs) proportionally without increasing at least one input (or decreasing at least one output). According to Koopmans's efficiency notion, a DMU is technically efficient if an increase in any output (or a decrease in any input) requires a decrease in at least another output (or an increase in at least another input).

From Figure 2.2, we can observe that DMU D is efficient in a Farrell (1957) sense but inefficient in a Koopmans (1951) sense, as it is possible to increase the amount of output y_2 from the level D to the level A, keeping the amount of y_1 produced with the same amount of input x. In a DEA assessment, although DMU D achieves an efficiency score equal to one, it is considered inefficient. The notion of efficiency adopted in a DEA assessment corresponds to the Koopmans's notion. The definition of efficiency in DEA is explained with more detail in section 2.2.3.

2.2.2 Production possibility set

In DEA, the production possibility set is derived from observed input-output combinations by making certain assumptions to the nature of the production possibility set. Consider a set of $j = 1, \ldots, n$ DMUs that use inputs $X \in \mathbb{R}_+^m$ to produce outputs $Y \in \mathbb{R}_+^s$. DMU j uses amount x_{ij} of input i $(i = 1, \ldots, m)$ to produce amount y_{rj} of output r $(r = 1, \ldots, s)$. The production possibility set (ϕ) contains all input-output feasible combinations corresponding to a certain production process. ϕ can be denoted as follows:

$$\phi = \{(X, Y) | \text{ Input vector } X \text{ can produce the output vector } Y\} \tag{2.5}$$

The assumptions postulated for the production possibility set can be defined as follows (Fried et al., 2008, p.255):

i) Convexity:
If $(X, Y) \in \phi$, and $(X^{'}, Y^{'}) \in \phi$ then $(\lambda(X, Y) + (1 - \lambda)(X^{'}, Y^{'})) \in \phi$ for any $\lambda \in [0, 1]$

ii) Monotonicity or strong free disposability of inputs and outputs:
If $(X, Y) \in \phi$ and $X^{'} \geq X$, then $(X^{'}, Y) \in \phi$
If $(X, Y) \in \phi$ and $Y^{'} \leq Y$, then $(X, Y^{'}) \in \phi$

iii) Inclusion of observations:
Each observed DMU $(X_{j_o}, Y_{j_o}) \in \phi$

iv) No output can be produced without some input:
If $Y \geq 0$, and $Y \neq 0$, then $(0, Y) \notin \phi$

v) Constant returns to scale:
If $(X, Y) \in \phi$ then $(\lambda X, \lambda Y) \in \phi$ for any $\lambda \geq 0$

vi) Minimum extrapolation:
ϕ is the intersection of all sets satisfying i) to v)

The previous assumptions allow to define a constant returns to scale (CRS) production possibility set, as follows:

$$\phi_{CRS} = \left\{ (X, Y) \in \mathbb{R}^{+} |\ X \geq \sum_{j=1}^{n} \lambda_j\ X_j,\ Y \leq \sum_{j=1}^{n} \lambda_j\ Y_j,\ \lambda_j \geq 0 \right\} \tag{2.6}$$

The CRS assumption implies that scaling up or down efficient input-output combinations is valid. However, in real life it is important to know how the technology behaves with changes in the scale of operation. This notion is captured by the returns to scale admitted by the technology. Returns to scale can be constant returns to scale (CRS) if output rises proportionally to input, increasing returns to scale (IRS) and decreasing returns to scale (DRS), such that output rises more than or less than proportionably to inputs, respectively. Dropping assumption v) from the previous set of assumptions allows to define a variable returns to scale (VRS) production possibility set, as follows:

$$\phi_{VRS} = \left\{ (X, Y) \in \mathbb{R}^{+} |\ X \geq \sum_{j=1}^{n} \lambda_j\ X_j,\ Y \leq \sum_{j=1}^{n} \lambda_j\ Y_j,\ \sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \geq 0 \right\} \tag{2.7}$$

The PPS under VRS (ϕ_{VRS}) differs from the CRS specification of the technology due to the additional convexity constraint $\sum_{j=1}^{n} \lambda_j = 1$. This constraint is in line with the convexity of the PPS assumed in ii) but is invalid under CRS, where any feasible and efficient input-output combination can be scaled up or down.

2.2.3 DEA models

DEA models with constant returns to scale

Model (2.4) is a fractional model but can be converted into a linear programming model through simple transformations, as shown in Charnes et al. (1978). The linearization of (2.4) can lead to an input oriented DEA model or to an output oriented DEA model, as shown in (2.8) and (2.9), respectively. Both formulations assume constant returns to scale.

$$
\begin{aligned}
& e_{j_o} = \max \sum_{r=1}^{s} u_r \; y_{rj_o} \\
& \text{subject to} \\
& \sum_{i=1}^{m} v_i \; x_{ij_o} = 1 \\
& \sum_{r=1}^{s} u_r \; y_{rj} - \sum_{i=1}^{m} v_i \; x_{ij} \leq 0, \;\; j = 1, \ldots, n \\
& v_i \geq \epsilon, \;\; i = 1, \ldots, m \\
& u_r \geq \epsilon, \;\; r = 1, \ldots, s
\end{aligned}
\tag{2.8}
$$

$$
\begin{aligned}
& h_{j_o} = \min \sum_{i=1}^{m} v_i \; x_{ij_o} \\
& \text{subject to} \\
& \sum_{r=1}^{s} u_r \; y_{rj_o} = 1 \\
& \sum_{r=1}^{s} u_r \; y_{rj} - \sum_{i=1}^{m} v_i \; x_{ij} \leq 0, \;\; j = 1, \ldots, n \\
& v_i \geq \epsilon, \;\; i = 1, \ldots, m \\
& u_r \geq \epsilon, \;\; r = 1, \ldots, s
\end{aligned}
\tag{2.9}
$$

For the input oriented perspective, the conversion into a linear programming model can be achieved by maximizing the numerator of the objective function in (2.4) and setting the denominator of the objective function equal to one as a restriction of the model. For the output oriented perspective, the linearization is done by minimizing the denominator of the objective function in (2.4) and setting the numerator of the objective function equal to one as a restriction of the model. The linearization of the restriction in (2.4) is trivial, and it becomes $\sum_{r=1}^{s} u_r \ y_{rj} - \sum_{i=1}^{m} v_i \ x_{ij} \leq 0$.

Models (2.8) and (2.9) are known as the "weights formulation" of the DEA model. The variables of the models are u_r and v_i, which correspond to the weights associated to the outputs and inputs, respectively. The input and output weights at the optimal solution can be used to indicate the relative importance of the inputs and outputs in determining the efficiency level of the DMU. However, these weights depend on the units of measurement of each output and input, so the "virtual" output ($u_r^* \ y_{rj_o}$) and "virtual" input ($v_i^* \ x_{ij_o}$) are used instead. The "virtual" values are normalized weights that do not depend on the scale of the variables, adding up to one for efficient DMUs both in terms of inputs and outputs.

In model (2.8) the relative efficiency score for DMU j_o is given by $e_{j_o}^*$, which reflects the proportion by which all inputs observed can be proportionally reduced, without reducing any output levels. In model (2.9), the relative efficiency score for DMU j_o is given by $1/h_{j_o}^*$, where $h_{j_o}^*$ corresponds to the proportion by which all outputs observed can be expanded proportionally, without requiring an increase to input level. In the case of constant returns to scale, the efficiency scores provided by the two models coincide, i.e. $e_{j_o}^* = 1/h_{j_o}^*$.

Using the duality of linear programming, we can derive equivalent forms for models (2.8) and (2.9), shown in models (2.10) and (2.11), respectively.

The dual form of the DEA models is referred to as the “envelopment formulation”. The dual models have a very intuitive interpretation that throws further light into the nature of the test of relative efficiency underlying a DEA analysis. The duality of linear programming states that the objective function value of the weight and envelopment problems is equal, corresponding to the efficiency score. In terms of by-products of the DEA assessment, the weights form provides information on the relative importance (weights) of the input and output variables, whereas the envelopment form provides information on peers and targets.

$$
\begin{aligned}
& e_{j_o} = \min\ \theta_o - \epsilon \left(\sum_{i=1}^{m} s_i + \sum_{r=1}^{s} s_r \right) \\
& \text{subject to} \\
& \theta_o\ x_{ij_o} - \sum_{j=1}^{n} \lambda_j\ x_{ij} - s_i = 0,\ i = 1, \ldots, m \\
& y_{rj_o} = \sum_{j=1}^{n} \lambda_j\ y_{rj} - s_r,\ r = 1, \ldots, s \\
& \lambda_j,\ s_i,\ s_r \geq 0,\ \forall_{j,i,r}
\end{aligned}
\tag{2.10}
$$

$$
\begin{aligned}
& h_{j_o} = \max\ \delta_o + \epsilon \left(\sum_{i=1}^{m} s_i + \sum_{r=1}^{s} s_r \right) \\
& \text{subject to} \\
& x_{ij_o} = \sum_{j=1}^{n} \lambda_j\ x_{ij} + s_i,\ i = 1, \ldots, m \\
& \delta_o\ y_{rj_o} - \sum_{j=1}^{n} \lambda_j\ y_{rj} + s_r = 0,\ r = 1, \ldots, s \\
& \lambda_j,\ s_i,\ s_r \geq 0,\ \forall_{j,i,r}
\end{aligned}
\tag{2.11}
$$

Models (2.10) and (2.11) seek to identify a comparator, i.e. a composite DMU corresponding to a linear combination of efficient DMUs $\left(\sum_{j=1}^{n} \lambda_j^* \, x_{ij},\right.$ $\left.\sum_{j=1}^{n} \lambda_j^* \, y_{rj}\right)$, with $i = 1, \ldots, m$ and $r = 1, \ldots, s$, that dominates DMU j_o in all input and output dimensions. For the inefficient DMUs, it is possible to obtain a set of targets to become efficient. The input and output targets for a DMU j_o are obtained, respectively, as shown in (2.12) and (2.13).

$$
\begin{aligned}
x_{ij_o}^{IT} &= \theta_o^* \, x_{ij_o} - s_i^* = \sum_{j=1}^{n} \lambda_j^* \, x_{ij}, \\
y_{rj_o}^{IT} &= y_{rj_o} + s_r^* = \sum_{j=1}^{n} \lambda_j^* \, y_{rj}.
\end{aligned}
\tag{2.12}
$$

$$
\begin{aligned}
x_{ij_o}^{OT} &= x_{ij_o} - s_i^* = \sum_{j=1}^{n} \lambda_j^* \, x_{ij}, \\
y_{rj_o}^{OT} &= \delta_o^* \, y_{rj_o} + s_r^* = \sum_{j=1}^{n} \lambda_j^* \, y_{rj}.
\end{aligned}
\tag{2.13}
$$

The targets correspond to a linear combination of the values observed in the peers. If $\lambda_j^* > 0$, then the corresponding DMU j is efficient and it is a peer to DMU j_o under assessment. Additional information obtained from these models relates to the slack variables (s_i and s_r). These indicate the extent to which individual inputs or outputs could be improved beyond the radial expansion corresponding to the efficiency score. If $s_i^* > 0$ or $s_r^* > 0$ for some input i or output r, the DMU j_o is projected on an inefficient segment of the frontier of the PPS. This distinguishes between Farrell and Koopmans efficiency notions. In Farrell's sense a DMU j_o is efficient if it has a radial efficiency score equals 1. In Koopmans's sense, a DMU j_o is efficient if it has a radial efficiency score equals 1, and no positive slack values, i.e. $s_i^* = s_r^* = 0, \;\; \forall_{i,r}$.

DEA models with variable returns to scale

The returns to scale is a characteristic of the frontier of the PPS. Returns to scale measure the response of output to equal proportional changes in all inputs. Banker et al. (1984) generalized the concept of returns to scale to the multiple input-output case, which can be expressed as:

- A DMU exhibits constant returns to scale (CRS) if a proportional increase (decrease) in the inputs causes an equal proportional increase (decrease) in the outputs.
- A DMU exhibits increasing returns to scale (IRS) if a proportional increase (decrease) in the inputs causes a greater than proportional increase (decrease) in the outputs.
- A DMU exhibits decreasing returns to scale (DRS) if a proportional increase (decrease) in the inputs causes a less than proportional increase (decrease) in the outputs.

The differences between an assessment under CRS and under VRS are illustrated in Figure 2.3, for a single input-output case.

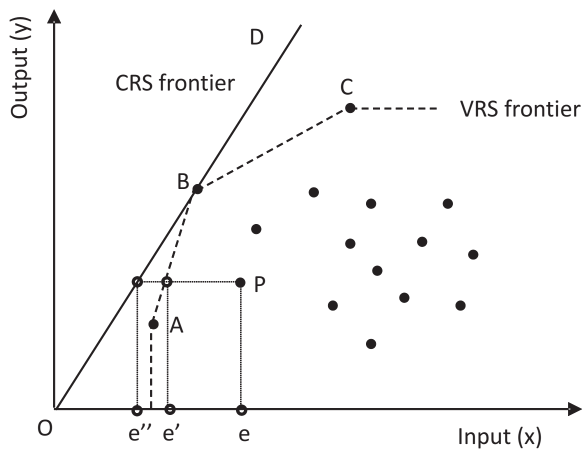

Figure 2.3: CRS and VRS frontiers

Under the assumption of CRS, DMU B can be extrapolated to points on the ray OD, such that the change in the input level causes an equally proportional change to the output level. The CRS frontier is defined by the ray OD.

If the scale extrapolation assumption used in the construction of the CRS frontier is not allowed, the frontier must be based on the observed performance of the DMUs given their scale of operation. The efficient frontier in Figure 2.3 is redefined as the segments between A, B, and C. This frontier allows for VRS and is made of convex combinations of the extreme points lying on the production surface.

Banker et al. (1984) extended the original DEA models of Charnes et al. (1978) to enable the estimation of efficiency under VRS. The corresponding "weights formulation" of VRS models with input and output orientations are shown in (2.14) and (2.15), respectively.

$$
\begin{aligned}
&\hat{e}_{j_o} = \max \sum_{r=1}^{s} u_r \; y_{rj_o} + \omega \\
&\text{subject to} \\
&\sum_{i=1}^{m} v_i \; x_{ij_o} = 1 \\
&\sum_{r=1}^{s} u_r \; y_{rj} - \sum_{i=1}^{m} v_i \; x_{ij} + \omega \leq 0, \; j = 1, \ldots, n \\
&v_i \geq \epsilon, \; i = 1, \ldots, m \\
&u_r \geq \epsilon, \; r = 1, \ldots, s \\
&\omega \in \mathbb{R}
\end{aligned}
\qquad (2.14)
$$

$$
\begin{aligned}
&\hat{h}_{j_o} = \min \sum_{i=1}^{m} v_i \ x_{ij_o} + \varpi \\
&\text{subject to} \\
&\sum_{r=1}^{s} u_r \ y_{rj_o} = 1 \\
&-\sum_{r=1}^{s} u_r \ y_{rj} + \sum_{i=1}^{m} v_i \ x_{ij} + \varpi \geq 0, \ j = 1, \ldots, n \\
&v_i \geq \epsilon, \ i = 1, \ldots, m \\
&u_r \geq \epsilon, \ r = 1, \ldots, s \\
&\varpi \in \mathbb{R}
\end{aligned}
\tag{2.15}
$$

Under VRS, the efficiency obtained is called pure technical efficiency (PTE), as opposed to the technical efficiency (TE) obtained under CRS. The pure technical efficiency of DMU j_o is given by $\hat{e}^*_{j_o}$ in model (2.14), and by $1/\hat{h}^*_{j_o}$ in model (2.15). Under VRS, the orientation of the assessment affects the segment of the projection and the resulting efficiency score may not be the same. For inefficient DMUs, it may occur $\hat{e}^*_{j_o} \neq 1/\hat{h}^*_{j_o}$, although the subset of efficient DMUs is the same independently of the orientation.

The dual of (2.14) and (2.15) correspond to the DEA "envelopment formulation" under VRS, as presented in (2.16) and (2.17), respectively. Note that to convert a CRS model into a VRS model, in the weights form, it requires an additional variable (ω or ϖ), whereas in the envelopment form it requires an additional constraint ($\sum_{j=1}^{n} \lambda_j = 1$). This convexity constraint is used under VRS to prevent any interpolation point constructed from the observed DMUs from being scaled up or down to form a referent point for efficiency measurement.

$$\begin{aligned}
&\hat{e}_{j_o} = \min\ \hat{\theta}_o - \epsilon\left(\sum_{i=1}^{m} s_i + \sum_{r=1}^{s} s_r\right)\\
&\text{subject to}\\
&\hat{\theta}_o\ x_{ij_o} - \sum_{j=1}^{n} \lambda_j\ x_{ij} - s_i = 0, i = 1, \ldots, m\\
&y_{rj_o} = \sum_{j=1}^{n} \lambda_j\ y_{rj} - s_r,\ r = 1, \ldots, s\\
&\sum_{j=1}^{n} \lambda_j = 1\\
&\lambda_j,\ s_i,\ s_r \geq 0,\ \forall_{j,i,r}
\end{aligned} \tag{2.16}$$

$$\begin{aligned}
&\hat{h}_{j_o} = \max\ \hat{\delta}_o + \epsilon\left(\sum_{i=1}^{m} s_i + \sum_{r=1}^{s} s_r\right)\\
&\text{subject to}\\
&x_{ij_o} = \sum_{j=1}^{n} \lambda_j\ x_{ij} + s_i, i = 1, \ldots, m\\
&\hat{\delta}_o\ y_{rj_o} - \sum_{j=1}^{n} \lambda_j\ y_{rj} + s_r = 0,\ r = 1, \ldots, s\\
&\sum_{j=1}^{n} \lambda_j = 1\\
&\lambda_j,\ s_i,\ s_r \geq 0,\ \forall_{j,i,r}
\end{aligned} \tag{2.17}$$

In the case of not knowing *a-priori* if the production technology exhibits CRS or VRS, it is possible to use hypothesis tests regarding returns to scale, such that the VRS model may only be used when returns to scale is confirmed. Banker (1996) was the first to use hypothesis tests in a DEA assessment. Later, Simar and Wilson (2002) proposed the use of bootstrapping to yield the appropriate critical values for the test statistics.

Comparing the distance of CRS and VRS frontiers at the scale size of the DMU it is possible to define a measure of scale efficiency (SE). For instance, in Figure 2.3, with an input orientation, the pure technical efficiency of DMU P in relation to the VRS frontier is $PTE = oe'/oe$, and its technical efficiency calculated in relation to the CRS frontier is $TE = oe''/oe$. The scale efficiency of DMU P is oe''/oe', which is equal to TE/PTE. The global measure of efficiency (measured in relation to the CRS frontier) is therefore a composite of pure technical efficiency and scale efficiency. With reference to Figure 2.3, we have $TE = oe''/oe = oe'/oe \times oe''/oe'$, which is equivalent to $TE = PTE \times SE$. The larger the divergence between VRS and CRS efficiency scores, the lower the value of scale efficiency, meaning that the impact of scale size on productivity is considerable.

2.2.4 DEA with weight restrictions

The DEA model is based on the assumption that each DMU should freely select weights for the inputs and outputs in order to ensure its efficiency is assessed in the best possible light. This reinforces certainty about inefficiency as inefficient units could not find a weighting scheme that conveys a 100% efficiency score. However, it may raise doubts about efficiency classifications as some units may appear efficient just because most inputs and/or outputs are assigned a very low weight (equal to ε). In practice, this means that these factors are in fact ignored in the efficiency assessment. It is possible to overcome this limitation by defining the range of values that the weights can take. This requires including weight restrictions in the DEA model.

Weight restrictions in a DEA model allow to respond to other problems raised in the literature, such as incorporating prior knowledge on the value of individual inputs or outputs, and improving discrimination between efficient DMUs. Weight restrictions have attracted considerable attention in the

DEA literature (see Allen et al., 1997 for a review). There are various types of weight restrictions that can be included in DEA models to explicitly link either input weights or output weights, which are presented next. When using weight restrictions, it is important to know the appropriate values for the parameters in the restrictions. A number of methods can be used to support the estimation of bounds to the weights, including, for example, the use of expert opinion, input/output price information, or marginal rates of substitution/transformation.

Absolute weight restrictions, first introduced by Dyson and Thanassoulis (1988), restrict weights to vary within a specific range. Such restrictions associated to output weights u_r and input weights v_i assume the form shown in (2.18) and (2.19), respectively. ρ_r and η_r (δ_i and τ_i) represent the bounds that the output weight (input weight) can take.

$$\rho_r \leq u_r \leq \eta_r \qquad (2.18) \qquad\qquad \delta_i \leq v_i \leq \tau_i \qquad (2.19)$$

These restrictions are used to prevent the inputs or outputs from being over or under emphasized in the assessment. The use of absolute weight restrictions has some practical difficulties. One is associated with the meaning of the bounds as weights are generally significant only on a relative basis, i.e. only the ratios of weights incorporate information concerning the marginal rates of transformation (between outputs) or substitution (between inputs). There are just a few cases where absolute bounds may have a clear interpretation. One of these examples is the single input multi-output case as presented in Dyson and Thanassoulis (1988), where output weights are interpreted as the level of input the DEA model assigns per unit of output. Other difficulty is that absolute weight restrictions may render the DEA models infeasible, and may not maximize the relative efficiency of the assessed DMU (for further details see Podinovski, 2001). Another concern associated to the use of absolute weight restrictions in a DEA model is that switching from an input to an output orientation produces different relative efficiency scores, even under CRS, and hence the bounds need to be set according to the orientation of the model. Due to the difficulties of implementing absolute weight restrictions, its use is limited to a few relatively simple cases (Dyson et al., 2001).

Assurance regions type I, first introduced by Thompson et al. (1986), model the relationship between output weights or input weights. Such restrictions, expressed in terms of output weights u_r and u_{r+1} and input weights v_i and v_{i+1}, assume the form shown in (2.20) and (2.21), respectively. θ_r and ζ_r (α_i and β_i) correspond to the bounds that the ratio of output weights (input weights) can take.

$$\theta_r \leq \frac{u_r}{u_{r+1}} \leq \zeta_r \qquad (2.20)$$

$$\alpha_i \leq \frac{v_i}{v_{i+1}} \leq \beta_i \qquad (2.21)$$

Assurance regions type I restrictions are used to incorporate in the analysis information concerning marginal rates of substitution/transformation between the inputs or outputs. As these restrictions impose the ratio of weights to be within a certain range, the bounds can be intuitively interpreted. It should be noted that the bounds are dependent on the scaling of the outputs and inputs, i.e. they are sensitive to the units of measurement of the related factors. Under CRS, assurance regions type I provide the same efficiency scores independently of the model orientation. This is because these restrictions work by changing the frontier, and efficiency is measured in relation to the modified frontier. This is demonstrated with an illustrative example in section 5.2.3.

Virtual weight restrictions, first introduced by Wong and Beasley (1990), are weight restriction applied to virtual inputs and outputs. These restrictions assume the form shown in (2.22) where the proportion of the total virtual output of DMU j devoted to output r, i.e. the importance attached to that output in percentual terms, is restricted to lie between ϕ_r and ψ_r. ϕ_r and ψ_r should be between 0 and 1.

$$\phi_r \leq \frac{u_r y_{rj}}{\sum\limits_{r=1}^{s} u_r y_{rj}} \leq \psi_r \tag{2.22}$$

$\sum\limits_{r=1}^{s} u_r y_{rj}$ corresponds to the total virtual output of DMU j, and is included in the denominator as a standardization way to facilitate the assignment of values of the bounds. As mentioned in section 2.2.3, input and output virtual weights indicate the relative importance of the inputs and outputs in determining the efficiency level of a DMU. In contrast, the weights u_r and v_i are scale dependent and should be interpreted carefully as a larger or smaller weight does not necessarily mean that a high or low importance is attached to a given input or output. One of the limitations of virtual restrictions is that they represent indirectly absolute bounds on the weights, so they are sensitive to model orientation. The other problem related to virtual restrictions is that they are DMU specific, and by adding these restriction to all DMUs and to several input/output variables, the model becomes computationally expensive. Wong and Beasley (1990) also suggest some modifications for implementing restrictions on virtual values trying to simplify the model. The modifications include adding restrictions only for the DMU under assessment or using constraints for an "average DMU".

2.2.5 Bootstrapping in DEA

The evaluation of performance in DEA involves the estimation of a frontier assuming that no random factors affect its construction. To overcome this limitation of the performance assessment, the DEA evaluation can be complemented with bootstrapping to ensure a robust evaluation of performance.

Bootstrapping, first introduced by Efron (1979), is a data-based simulation method for statistical inference. Bootstrapping is based on the idea that when little or nothing is known of the underlying data generating process for a sample of observations, the data generating process can be estimated using the original sample to generate a bootstrapped sample from which the parameters of interest can be derived. The process involves using the original sample to construct an empirical distribution of the variables of interest through the repeated sampling of the original data set. Using bootstrapping it can be determined the appropriate confidence intervals for the scores estimated, and whether these scores are statistically significant.

The basic idea behind the bootstrapping method (naive bootstrapping) can be explained as follow. Consider that we draw a sample of a given variable (output) $S = (y_1, \ldots, y_n)$, where n is the number of DMUs from a population $P = (y_1, \ldots, y_N)$, being N much larger than n. Suppose that we are interested in some statistic $T = t(S)$ as an estimate of the corresponding population parameter $Q = t(P)$. Suppose that we proceed to draw a sample of size n from the elements of S, sampling with replacement. Call the resulting bootstrap sample $S^{1*} = (y_1^{1*}, \ldots, y_n^{1*})$. A sampling procedure with replacement is needed as otherwise we would simply reproduce the original sample S. In fact, we are treating the sample S as an estimate of the population P, i.e., each element y_i of S is selected for the bootstrap sample with probability $1/n$, mimicking the original selection of the sample S from the population P. We repeat this procedure a large number of times,

B, generating many pseudo-samples. The *bth* bootstrap sample is denoted $S^{b*} = (y_1^{b*}, \ldots, y_n^{b*})$. We compute the statistic T for each of the pseudo-samples, that is $T^{b*} = t(S^{b*})$. The distribution of T^{b*} around the original estimate T is analogous to the sampling distribution of the estimator T around the population parameter Q. Concerning the number of bootstrap samples, we can enumerate *all* bootstrap samples of size n considering n^n bootstrap samples. However, it can be almost impossible to do this in practice unless n is tiny. To avoid this, we draw at random a large number of bootstrap samples. The results presented by Efron and Tibshirani (1993) suggest that using 2000 bootstrap replications provides accurate results.

Simar and Wilson (1998) proposed the use of bootstrapping in the context of DEA efficiency assessments, in order to obtain unbiased efficiency estimates and to calculate confidence intervals for the efficiency point estimates. Efficiency scores computed by DEA models are truncated, with an upper value equal to one, and there may exist several estimates close to unity. Consequently, resampling directly from the original data (naive bootstrap) provides a poor estimate of the data generating process. The common approach is to nonparametrically estimate the original densities of the performance scores using kernel smoothing methods, combined with a reflection method (Silverman, 1986). Simar and Wilson (1998) proposed the smoothed bootstrap method which deals with all the particular features of DEA-based efficiency scores when mimicking the data generating process. The bootstrapping process can be summarized as follows.

1. Compute the performance estimates $\widehat{E}_j$ for each DMU $j = 1, \ldots, n$, using a DEA model.

2. Use kernel density estimation and the reflection method (smooth bootstrap) to generate a random sample of size n from $\{\widehat{E}_j, j = 1, \ldots, n\}$, resulting in $\{E^*_{jb}, j = 1, \ldots, n\}$.

3. Generate a pseudo data set $\{(x^*_{jb}, y^*_{jb}),\ j = 1, \ldots, n\}$ to form the reference bootstrap technology, where $y^*_{jb} = (\widehat{E}_j / E^*_{jb}) \times y_j$ and $x^*_{jb} = x_j$, $j = 1, \ldots, n$, for an output oriented perspective (as used in this thesis).

4. Compute the bootstrap estimate of efficiency $\widehat{E}^*_{jb}$ of $\widehat{E}_j$ for each $j = 1, \ldots, n$, using a DEA model.

5. Repeat steps 2-4 B times (B=2000) to obtain a set of estimates $\{\widehat{E}^*_{jb}, b = 1, \ldots, B\}$.

Having the bootstrap values computed, we can obtain the bias of $\widehat{E}_j$ as

$$\widehat{bias}_B(\widehat{E}_j) = B^{-1} \sum_{b=1}^{B} \widehat{E}^*_{jb} - \widehat{E}_j, \tag{2.23}$$

the bias-corrected estimates of E_j as

$$\widehat{\widehat{E_j}} = \widehat{E}_j - \widehat{bias}_B(\widehat{E}_j) = 2\widehat{E}_j - B^{-1} \sum_{b=1}^{B} \widehat{E}^*_{jb}. \tag{2.24}$$

The construction of confidence intervals for E_j involves the following steps:

- sort the values $(\widehat{E}^*_{jb} - \widehat{E}_j)$ for $b = 1, \ldots, B$ in increasing order, and delete $\left(\frac{\alpha}{2} \times 100\right)\%$ of the elements at either end of the sorted array;

- set $-\widehat{b}^*_\alpha$ and $-\widehat{a}^*_\alpha (\widehat{a}^*_\alpha \leqslant \widehat{b}^*_\alpha)$, equal to the endpoints of the sorted array. The estimated $(1 - \alpha)\%$ confidence interval is obtained as

$$\widehat{E}_j + \widehat{a}^*_\alpha \leq E_j \leq \widehat{E}_j + \widehat{b}^*_\alpha \tag{2.25}$$

2.3 Malmquist productivity index

2.3.1 Introduction to productivity change

Firstly, it is important to distinguish between efficiency and productivity concepts. Efficiency is a relative measure, and it is defined by comparing the input and output of a unit with those of the best performing units from its peers. Productivity is an absolute concept. The productivity of a $unit_j$ is defined as the amount of output produced per unit of input used. In the case of a single input and output, it is defined as:

$$\text{Productivity} = \frac{output}{input} \tag{2.26}$$

The standard approach within the non-parametric literature to evaluate productivity over time is the Malmquist productivity index (MI). The MI was first introduced by Caves et al. (1982), and then developed by Fare et al. (1994b), who applied DEA to estimate the index. The use of Malmquist index to measure productivity change over time has many interesting features. For instance, it can be based on multi input-output frontier representations of the production technology, and does not require price information. Another important advantage is that the index can be decomposed into efficiency change and technological change, providing insights into the root sources of productivity change.

Caves et al. (1982) defined an output oriented and an input oriented productivity index. The output oriented version of the index is presented herein as it is the index used in the empirical part of this thesis. Consider n DMUs in time period t that use inputs $x^t \in R_+^m$ to produce outputs $y^t \in R_+^s$, and in time period $t+1$ use inputs $x^{t+1} \in R_+^m$ to produce outputs $y^{t+1} \in R_+^s$. According to Caves et al. (1982), the Malmquist index allows to compare the performance of DMUs between time period t (x^t, y^t) and time period

$t+1$ (x^{t+1}, y^{t+1}). In period t the production technology (ϕ^t) can be defined as shown in (2.27). It consists of all input-output combinations that are technically feasible for a certain production process.

$$\phi^t = \{(x^t, y^t) : x^t \text{ can produce } y^t\} \tag{2.27}$$

The output distance function for DMU j in relation to the technology (ϕ^t) can be defined as shown in (2.28)(Shepard, 1953):

$$D_o^t(x^t, y^t) = \min\left\{\theta : \left(x^t, \frac{y^t}{\theta}\right) \in \phi^t\right\},\ \theta \leq 1 \tag{2.28}$$

The output distance function for each DMU j is the reciprocal to the maximal feasible expansion of output y^t producible from input x^t. This means that it corresponds to the DEA efficiency score of each DMU in period t, i.e. $D_o^t(x^t, y^t) \leq 1$.

The output oriented Malmquist index for each DMU j relative to technology t is the ratio between the output distance function estimated for DMU j in period $t+1$, (x^{t+1}, y^{t+1}), relative to the technology in period t, and the output distance function estimated for DMU j in period t, (x^t, y^t), relative to the technology in t, as shown in (2.29):

$$M_o^t = \frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \tag{2.29}$$

Similarly, the output oriented Malmquist index for each DMU j relative to technology $t+1$ is defined as:

$$M_o^{t+1} = \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^{t+1}(x^t, y^t)} \tag{2.30}$$

Fare et al. (1994b) defined an output oriented productivity index as the geometric mean of the two indexes using as reference the technology at time periods t and $t+1$, yielding the following Malmquist index:

$$MI^{t+1,t} = \left(\frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \cdot \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^{t+1}(x^t, y^t)} \right)^{\frac{1}{2}} \tag{2.31}$$

Another major achievement of Fare et al. (1994b) was to show how to decompose the Malmquist index into an index of technical efficiency change and an index of technological change. These components are obtained by rewriting the index as follows:

$$MI^{t+1,t} = \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \cdot \left(\frac{D_o^t(x^t, y^t)}{D_o^{t+1}(x^t, y^t)} \cdot \frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^{t+1}(x^{t+1}, y^{t+1})} \right)^{\frac{1}{2}} = EC \cdot TC \tag{2.32}$$

The first component (EC), outside the square bracket, reflects the relative change in efficiency between periods t and $t+1$, i.e. measures how the DMUs have behaved in catching up with the others on the frontier. The second component (TC), corresponding to the geometric mean of the two ratios in square brackets, reflects the relative distance between the frontiers of the PPS in t and $t+1$, i.e. captures the distance between the frontiers of the two periods evaluated at the input-output levels at t (x^t, y^t) and at $t+1$ (x^{t+1}, y^{t+1}). The values of $MI^{t+1,t}$ may be greater, equal or smaller than one, depending on whether productivity growth, stagnation or decline occurred between periods t and $t+1$. A similar interpretation applies to EC and TC components.

Note that productivity growth may involve technological regress (if gains in efficiency dominate the regress in the frontier of the PPS from t to $t+1$) or a decline in efficiency (if technological progress in the frontier of the PPS from t to $t+1$ dominates the loss in the efficiency). Similar possibilities hold for the case of productivity decline.

The output oriented Malmquist index requires the estimation of within-period and mixed-period distance functions, which can be obtained using DEA models. Whereas the single period distance function is always less than or equal to one, in the mixed-period assessments required for the estimation of the Malmquist index, the value of the output distance function may be smaller, equal or greater than unity. This is because the input-output combination observed in one period may not be feasible within the technology in another period.

The basic ideas behind the calculation of the MI are explained using an illustrative example, with two outputs (y_1) and (y_2) and one input (x). Figure 2.4 illustrates a production possibility set at time periods t and $t+1$, with an output oriented perspective. The calculation of productivity change for DMU C, represented by point a on time period t and point f on time period $t+1$ will be explained in detail.

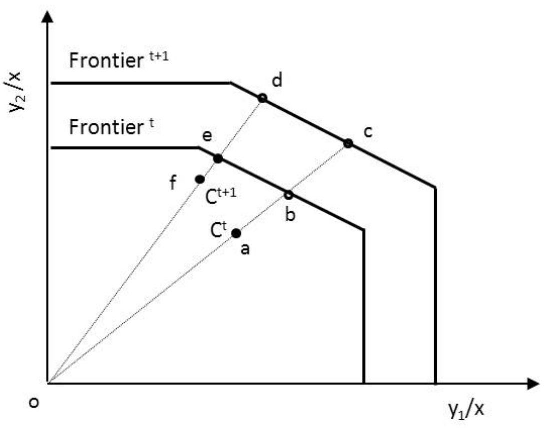

Figure 2.4: Malmquist output oriented productivity index

In the example, the efficiency score of DMU C in period t measured in relation to the technology in period t, corresponds to the ratio oa/ob. The efficiency score of DMU C in period $t+1$ measured in relation to the technology in period $t+1$, corresponds to the ratio of/od. Hence, the change in efficiency between period t and $t+1$ is measured as follows:

$$EC = \left(\frac{of/od}{oa/ob}\right) \tag{2.33}$$

The distance between the two frontiers measured along ray C^t can be expressed by $(oa/ob)/(oa/oc)$. The numerator of this ratio (oa/ob) represents the efficiency of DMU C in period t in relation to the frontier in period t, and the denominator (oa/oc) corresponds to the efficiency of DMU C in period t, measured in relation to the frontier in period $t+1$. This term (oa/oc) is called a mix-period efficiency score. Similarly, the distance between the two frontiers measured along ray C^{t+1} can be obtained by $(of/oe)/(of/od)$. The geometric mean of the two ratios defines the change in technology between period t and $t+1$ as follows:

$$TC = \left(\frac{oa/ob}{oa/oc} \cdot \frac{of/oe}{of/od}\right)^{\frac{1}{2}} = \left(\frac{oc}{ob} \cdot \frac{od}{oe}\right)^{\frac{1}{2}} \tag{2.34}$$

Values of TC greater than one signal technological improvement, which is the case in this illustrative example, as the frontier in $t+1$ expanded the PPS of period t.

The Malmquist productivity index is then the product of efficiency change and technological change indexes, as shown in 2.35.

$$MI^{t+1,t} = \left(\frac{of/od}{oa/ob}\right) \cdot \left(\frac{oc}{ob} \cdot \frac{od}{oe}\right)^{\frac{1}{2}} \tag{2.35}$$

2.3.2 Bootstrapping the Malmquist index

Bootstrapping in Malmquist indexes is used to test if changes in productivity, efficiency and technology are significant, i.e whether the estimates indicate a real change in performance or if they are a result of sampling noise (Simar and Wilson, 1999). Bootstrapping in Malmquist index is similar to bootstrapping in DEA. However the autocorrelation in the panel data has to be taken into account, as observations concerning the same DMU may be correlated across time periods. Simar and Wilson (1999) addressed this problem by extending the smooth bootstrap, which required estimating the joint density of inputs and outputs over two time periods. This enables estimating confidence intervals for the Malmquist index and its components.

The confidence intervals allow to verify whether the MI and its components are significantly different from one. If the interval contains the value one, we cannot infer that significant changes occurred in the performance of the DMU under assessment. Conversely, if the lower and upper bounds are smaller (or greater) than one, it is an indication that productivity declined (or improved).

The procedure to bootstrap the Malmquist index can be summarized in the following steps:

1. Compute the Malmquist index $\widehat{MI}_j(t_1, t_2)$ for each DMU $j = 1, \ldots, n$, using DEA models.
2. Use bivariate kernel density estimation and the reflection method (smooth bootstrap accounting for time dependence) to generate a pseudo data set $\{(x^*_{jt}, y^*_{jt}),\ j = 1, \ldots, n,\ t = 1, 2\}$.
3. Compute the bootstrap estimate of Malmquist index $\widehat{MI}^*_{jb}(t_1, t_2)$ applying the original estimators to pseudo data set derived from step 2.

4. Repeat steps 2 and 3 B times (B = 2000) to obtain a set of estimates $\{\widehat{MI}^*_{j1}(t_1, t_2), \ldots, \widehat{MI}^*_{jB}(t_1, t_2)\}$.

Having the Malmquist bootstrap values computed, we can obtain the bias of $\widehat{MI}_j(t_1, t_2)$ as

$$\widehat{bias}_B(\widehat{MI}_j(t_1, t_2)) = B^{-1}\sum_{b=1}^{B}\widehat{MI}^*_{jb}(t_1, t_2) - \widehat{MI}_j(t_1, t_2), \tag{2.36}$$

the bias-corrected estimates of $MI_j(t_1, t_2)$ as

$$\widehat{\widehat{MI}}_j(t_1, t_2) = \widehat{MI}_j(t_1, t_2) - \widehat{bias}_B(\widehat{MI}_j(t_1, t_2)) = 2\widehat{MI}_j(t_1, t_2) - B^{-1}\sum_{b=1}^{B}\widehat{MI}^*_{jb}(t_1, t_2). \tag{2.37}$$

The construction of the confidence intervals for MI_j involves the following steps:

- sort the values $(\widehat{MI}^*_{jb}(t_1, t_2) - \widehat{MI}_j(t_1, t_2))$ for $b = 1, \ldots, B$ in increasing order, and delete $\left(\frac{\alpha}{2} \times 100\right)\%$ of the elements at either end of the sorted array;
- set $-\widehat{b}^*_\alpha$ and $-\widehat{a}^*_\alpha(\widehat{a}^*_\alpha \leqslant \widehat{b}^*_\alpha)$, equal to the endpoints of the sorted array. The estimated $(1-\alpha)\%$ confidence interval is obtained as

$$\widehat{MI}_j(t_1, t_2) + \widehat{a}^*_\alpha \leq MI_j(t_1, t_2) \leq \widehat{MI}_j(t_1, t_2) + \widehat{b}^*_\alpha \tag{2.38}$$

The procedures described can be repeated to obtain bootstrap estimates for efficiency change (EC) and technological change (TC).

2.4 Summary and conclusions

This chapter presented the techniques that are more relevant to achieve the objectives of this thesis. It emphasized the DEA technique, as it will be used throughout the thesis. In particular, the chapter reviewed the assumptions underlying the construction of the production possibility set, the standard DEA models and associated measures of efficiency (i.e. technical efficiency, pure technical efficiency, and scale efficiency), and the use of weight restrictions in DEA models.

This chapter also presented the Malmquist index, which is considered the standard approach to measure productivity change over time. In addition, the bootstrapping procedures to complement a DEA efficiency assessment and the estimation of the Malmquist index were described.

Chapter 3

The analysis of performance in the Construction Industry

3.1 Introduction

This chapter provides an overview of the literature concerning performance measurement in the construction industry. The objective of this review is to summarize the aims, methods and conclusions of existing studies in order to identify fruitful research directions.

Due to the increasingly fierce competitive environment in the construction industry, the performance assessment in the sector has attracted considerable attention over the past 15 years (see Bassioni et al. (2004), for a literature review). The literature on performance measurement in the construction industry has been concerned not only with the development of performance assessment systems for construction companies, but also with the development of specific systems aiming to support the selection of construction companies during the bidding process.

This chapter is structured as follows. Section 3.2 presents a brief state-of-the-art review on the performance assessment systems developed for the construction industry, describing with more detail the benchmarking tools frequently used in this sector, as well as more recent developments concerning the application of DEA in the CI. Section 3.3 includes a brief literature review on the performance assessment systems aiming to support the selection of companies during the bidding process. Section 3.4 summarizes and concludes.

3.2 Performance measurement in the CI

Traditionally, performance measurement in the construction industry relied on financial indicators, such as profitability or return on capital. A company would be evaluated comparing its financial indicators with the average value of the industry. However, construction companies became more complex organizations with a multivariate nature, and judging performance merely based on a financial diagnosis resulted in an inadequate assessment. In particular, the financial approach to evaluate performance has several drawbacks, such as: i) it only reflects the current achievements, ii) it has a short-term nature, iii) it fails to link current performance with future performance, iv) it hinders innovation, v) it has an internal focus, providing little information on competitors, suppliers, and clients.

During the 1990's, company-wide approaches to measure performance, including financial and nonfinancial indicators were introduced in the construction industry. Kaplan and Norton (1992) revolutionized the financial approach by developing the Balanced Scorecard. The Balanced Scorecard complements traditional financial measures by incorporating criteria that measures performance from three additional perspectives, namely customer satisfaction, internal business processes, and innovation and learning. The

Balanced Scorecard also puts an emphasis to link company long-term strategy with its short-term actions. Since then, many researchers have developed other performance measurement systems, including new philosophies and dimensions of performance. Among them, the following are frequently mentioned in the literature: the strategic measurement analysis and reporting technique (SMART) system (Lynch and Cross, 1991), and the Performance Prism (Neely et al., 2002). Other company-wide performance frameworks used by companies are based on quality management models. The most commonly used models are the excellence model developed by the European Foundation for Quality Management (EFQM) (www.efqm.org), the Malcolm Baldrige model from the National Institute of Standards and Technology (www.baldridge.com) in the United States, and the Deming Prize model developed by the Union of Japanese Scientists and Engineers (www.juse.or.jp). A few researchers have tailored these frameworks to the specific needs of the CI. For instance, Kagioglou et al. (2001) developed a conceptual performance measurement framework based on the Balanced Scorecard, adding "project" and "supplier" perspectives. Beatham et al. (2002) reported different uses of the EFQM excellence model applied to the construction industry, and Bassioni et al. (2005) proposed a framework that combined the Balanced Scorecard and EFQM excellence model.

After the publication of the seminal reports of Latham (1994) and Egan (1998), construction companies have been mostly concerned with benchmarking systems based on performance indicators. These systems are usually available in the internet, and enable collecting data and producing real-time results concerning performance levels. In several countries, benchmarking systems were specially developed for the construction industry. The first benchmarking initiative was launched in the United Kingdom, called "Key Performance Indicators", and it is currently lead by the Construction Excellence organization (www.constructingexcellence.org.uk). Nowadays, Key

Performance Indicators are a tool used by many construction organizations worldwide.

Other relevant construction industry web benchmarking platforms appeared later. Costa et al. (2006) describe the scope of four well known benchmarking programs focused on construction performance measures, carried out in Brazil, Chile, United Kingdom and United States. As explained by Costa et al. (2006), these benchmarking programs typically aim to i) offer guidance for performance measurement; ii) provide benchmarks that can be used by individual companies to establish business goals and objectives; iii) identify and disseminate the best practices in the industry through reports and benchmarking networks. The authors also discussed the lessons learned and improvement opportunities that were identified in the design and implementation of the benchmarking programs of these countries.

The benchmarking program developed by the Construction Industry Institute in the United States was described with particular detail by Lee et al. (2005). The web system, called Benchmarking and Metrics (BM&M), enables collecting data and producing results concerning performance levels and best practices. Additionally, Ramirez et al. (2004) describes the benchmarking system that was established in the Chilean CI, which incorporates qualitative management aspects and performance indicators. The authors report the results obtained from the initial application of the management evaluation system, including different analysis to determine trends in the sector and to establish correlations between qualitative and quantitative aspects.

In Portugal, a web benchmarking platform for the CI, called icBench, was also developed as described in Costa et al. (2007). The icBench is a web platform to evaluate company performance through a set of KPIs, and allows continuous data collection and real time generation of outcomes. The platform was designed to provide two types of functionalities. One is to act as a self-evaluation tool for companies, as it supports a data management system updated by the companies every year. The other is to allow external benchmarking comparisons. Construction companies have to fill periodically on-line questionnaires to provide data on organizational and operational performance indicators. After submitting the questionnaires, the system automatically generates the benchmarking results that become available to companies. The outcomes are presented using ranking curves that show the benchmark score for each KPI of each company, and using radar charts that give a picture of the company overall performance. The icBench platform aims to be a friendly tool for providing guidance to construction companies, ensuring that managers are identifying and targeting the right problems. The platform was launched in 2006 and was developed at Faculty of Engineering of the University of Porto. It is currently sponsored by the regulatory board of Portuguese Construction and Real Estate (INCI).

More recently, the potential for using frontier methods to analyze performance in the CI has started to be explored. In particular, the literature has described successful applications of the DEA technique to the CI. El-Mashaleh et al. (2007) considered DEA a tool that can offer significant improvements over web benchmarking systems for the CI. According to the authors, the existing benchmarking systems have limitations in their ability to guide the industry towards more efficient and effective performance. Using the DEA feature of being able to summarize several firm performance metrics into a single score, they developed DEA models that allow construction firms to be evaluated on a company-wide basis. El-Mashaleh et al. (2006) used DEA to obtain a single efficiency score that reflects overall firm performance based on the aggregation of several performance indicators. This allowed analyzing the impact of Information Technology on construction firm performance. The study demonstrated that an increase in Information Technology utilization produces an increase in overall firm efficiency. Pilateris and McCabe (2003) developed a model based on DEA technique to assess contractor financial performance and to identify a set of efficient companies to be considered as benchmarks for the Canadian construction industry. The results were obtained by construction activity profile (buildings, heavy civil and specialty trade), and for five regions across Canada.

The studies previously described focus on the performance of individual companies, providing insights concerning the strengths, weaknesses and targets for improvement. Other studies focus on the assessment of the CI sector of particular countries. For instance, You and Zi (2007) analyzed the cost, allocative and technical efficiency of the Korean CI in the period 1996 to 2000, which includes the period that the country faced an economic crisis. The results showed that efficiency decreased significantly during the period analyzed. Xue et al. (2008) used the Malmquist index to measure the productivity change of construction companies from different Chinese regions. The results of the analysis indicated that Chinese companies experienced a continuous productivity improvement between 1997 and 2003, with the exception of year 2002.

Further research concerning the assessment of performance in the CI is in need. In particular, it would be important to extend the analysis to identify the factors that promote better performance levels in the CI. This topic has attracted attention in areas such as manufacturing (e.g. Narasimhan et al., 2005; Sousa and Voss, 2002), transportation (e.g. Fung et al., 2008; Odeck, 2006, 2008; Barros and Peypoch, 2009; Chen et al., 2008), banking (e.g. Mukherjee et al., 2001), agriculture (e.g. Latruffe et al., 2008; Chen et al., 2008), or iron and steel industry (e.g. Ma et al., 2002). The advanced status of the state-of-the-art in other sectors served as a reference for the research of this thesis. The considerable number of scientific publications available in different sectors enabled transferring the methodologies used to the construction industry context.

Comparisons of the CI sectors of different countries have not been explored to date. As the construction markets of most countries are becoming increasingly integrated, it would be important to broaden the scope of the performance assessment to promote learning from the best practices observe in other companies from different regions. Cross-country comparisons were only conducted in other activity sectors. For instance, Rao et al. (2003) used the “metafrontier” concept of Hayami and Ruttan (1971) to compare the efficiency levels of the agriculture sector in four different regions of the world (Africa, America, Asia and Europe) from 1986 to 1990. In banking, Pastor et al. (1997) compared the efficiency and the productivity differences between European and United States banking systems for the year of 1992 using DEA and Malmquist index, and Johnes et al. (2009) compared the efficiency of banks from the Gulf Cooperation Council located in different countries from 2004 to 2007 using financial ratio analysis and DEA. In electricity, Haney and Pollitt (2009) presented the results of an international survey of energy regulators in 40 countries from Europe, Australia, Asia and Latin America.

3.3 Specific performance measurement systems

This section presents a literature review related to the performance assessment systems developed specially to support the selection of companies during the bidding process. The bidding process in the construction industry occurs between the owner and the general contractor (GC) and between the GC and the subcontractors (SCs). The owner selects the appropriate GC to manage and carry the construction project, and the GC selects the SCs for undertaking specific tasks. As mentioned by Kumaraswamy and Matthews

(2000) and Arditi and Chotibhongs (2005), the literature concentrates more on the evaluation of GC rather than on the evaluation of SCs.

Models for evaluating GCs are known as prequalification models in the literature. A review on the prequalification models used in United States, Hong Kong, and Australia can be found in Palaneeswaran and Kumaraswamy (2001). In the literature, there are various prequalification models that differ in terms of the techniques used. They vary from simple rankings of GCs, based on arithmetic averages, or weighted averages (Russell and Skibniewski, 1990) to more complex quantitative techniques, such as cluster analysis (Holt, 1996), analytical hierarchy process (Abudayyeh et al., 2007; Topcu, 2004), multivariate statistical techniques (Wong et al., 2003), fuzzy set theory (Singh and Tiong, 2005), evidential reasoning (Sonmez et al., 2002), multicriteria utility theory (Hatush and Skitmore, 1998) or graph theory and matrix methods (Darvish et al., 2009).

Data Envelopment Analysis has also been recognized as a useful tool to support the selection of GC during the bidding process. There are two studies in the literature using the DEA technique to select GCs. The study conducted by McCabe et al. (2005) developed a contractor prequalification system using a DEA model combined with a methodology for determining a "practical frontier" of best contractors. The frontier specified can be used as a regional performance standard, which can help owners in the selection of GCs and provide guidance for improvement to company managers. The study developed by El-Mashaleh (2010) proposed a DEA model to guide owners in the bidding process. The author considered that the best bids constitute the DEA frontier, and correspond to the best candidates in the selection process. Another study worth mentioning is the one developed by Castro-Lacouture et al. (2007). Although the authors did not develop a system to support the selection of GCs, they presented a DEA-based tool for optimizing purchase decisions in construction e-marketplaces. This tool allows large communities of buyers and suppliers to meet and trade with each other.

Other studies in the literature focus on developing prequalification models with specific characteristics. For instance, Palaneeswaran and Kumaraswamy (2000) proposed a model for GC prequalification, specially developed for the analysis of public projects, and Abudayyeh et al. (2007) developed a methodology for submitting tenders in public projects using one of the following three types of contracting methods: design-build, cost-plus-time, and warranty. Alarcon and Mourgues (2002) proposed a contractor selection system including the contractor performance prediction as one of the criteria.

Concerning the performance assessment systems to support the selection of subcontractors, a few examples can be found in the literature. The studies differ in terms of the methods used to rank subcontractors. For instance, Albino and Garavelli (1998) proposed a neural network application to support SC rating. Ko et al. (2007) developed a subcontractor performance evaluation model, called SPEM, by employing an evolutionary fuzzy neural inference model to execute the evaluation process. Elazouni and Metwally (2000) developed a decision support system, called D-Sub, using linear programming and financial analysis techniques. The objective of the model was to minimize the total cost, analyzing the trade-off between using internal resources or subcontracting the work.

Other studies focus on improving the usability of existing systems to foster their application in practice. For instance, by making the user interface friendlier such as in Arslan et al. (2008), or by speeding up the subcontracting process using an integrated extensible markup language (XML), such as in Tserng and Lin (2002). Other studies incorporate particular features in the SC evaluation systems. For instance, Maturana et al. (2007) developed an on-site evaluation system based on lean principles and partnering practices.

The approaches proposed in the literature still have scope for improvement, as none of the existing systems can accommodate simultaneously the effective needs of the construction companies such as: i) decisions based on various perspectives; ii) bilateral evaluations to deeper understand contractor-supplier relationship; iii) flexibility to include the decision maker preferences in terms of the relative importance of the criteria.

3.4 Summary and conclusions

This chapter reviewed the literature on performance measurement in the CI, summarizing the main aims, models used, and conclusions of existing studies. The information gathered contributed to identify the issues that need further attention. These issues are addressed in the following chapters of this thesis by the development of innovative models and methodologies to evaluate performance in the CI. The relationship between the gaps identified and the chapters of this thesis is explained next.

It is fair to conclude that the performance assessment systems developed to support the selection of companies during the bidding process still have room for improvement. Chapter 4 addresses this issue by developing a performance assessment model based on several criteria, enabling bilateral evaluations between companies and incorporating the decision maker viewpoint when selecting the appropriate company for a given work.

Performance measurement in the construction industry typically involves the use of key performance indicators, usually available in web benchmarking platforms. However, no insights concerning organization overall performance and targets for performance improvement are available. Chapter 5 aims to fulfill this gap by proposing a methodology that uses Data Envelopment

Analysis as a method to complement the information provided by a set of key performance indicators.

The literature on performance measurement is scarce in studies using frontier methods to analyze the construction sector of particular countries. Chapter 6 addresses this issue by developing a comprehensive model to evaluate performance and innovation within the Portuguese construction industry, identifying the factors that promote excellence in the construction industry companies.

To the best of our knowledge, comparisons of the construction sector from different countries were not explored to date. Chapter 7 aims to fulfill this gap by developing a robust approach to evaluate the efficiency of construction companies worldwide, taking into account the company geographic location and company activity profile. In addition, Chapter 7 proposes an approach to analyze the general issue of the convergence in efficiency across regions.

Chapter 4

Design of a performance assessment system for the selection of contractors in construction industry e-marketplaces

4.1 Introduction

The developments in the construction industry (CI) have resulted in a significant increase of subcontracting practices. Nowadays, general contractors (GCs) are focused mainly on core activities, and do not participate in peripheral tasks associated with the completion of the construction project. This means that GCs employ a minimum workforce in construction projects, promoting the specialization of construction companies (Maturana et al., 2007). For selecting the subcontractors (SC) to undertake specialized works, some of the major duties of the GC are to prepare tender documents, evaluate bids, and award the contract to the bid winner. The owner follows the same sequential procedure to select the GC to manage and carry the construction project. The bidding process between the different agents corresponds to one of the most crucial phases that determines the success of a construction project.

Traditionally, the competitive bidding was decided based on a price criterion, neglecting aspects related to company quality. This could result in an inaccurate selection of companies, eliminating qualified companies and accepting others incapable of undertaking the job successfully. This would have serious repercussions in the construction projects in terms of poor quality works, delays in project duration, additional costs, and reworks. A discussion of the major problems associated with the traditional company selection methods can be found in Tserng and Lin (2002). To overcome these problems, the selection of the most appropriate company for a specific work should be based on a set of criteria, such as company technical capability and financial stability.

Recently, e-marketplaces have started to be developed in order to facilitate the bidding process. The traditional bidding process, from the beginning of the tender document preparation to the awarding of the bid, is now carried out in a paperless manner. The e-marketplaces enable to promote a collaborative working environment, provide transparency, and accelerate the bidding process. The use of e-marketplaces by construction companies is increasing exponentially worldwide. In Portugal, the code of public contracts, launched in 2008, transposes the European directives to the national context. This code obliges public works to be announced in an e-marketplace. With the growing number of proposals shared in e-marketplaces, decision making is increasingly more complex. It is essential to incorporate consistent performance assessment systems to facilitate the decision concerning the selection of the most appropriate company for a given work.

The purpose of this chapter is to present a performance assessment system designed with the purpose to evaluate construction companies based on several criteria. This system was intended to be integrated in e-marketplaces in order to comply with the technological advances in the CI. The system, named CIsea - Construction Industry system for efficiency assessment, aims to facilitate the selection of the best subcontractor among competitive bids and incorporates innovative features such as the bilateral evaluation between companies.

The development of CIsea started in a few meetings undertaken with members of Vortal (www.vortal.pt). Vortal is one of the leading Portuguese companies of e-marketplaces operating in various activity sectors (e.g. construction, health care, energy and utilities). In particular, the system was designed to be integrated in the Vortal e-marketplace for the Portuguese construction industry. The design of CIsea was followed by a team from Vortal to ensure that the system accomplished the features required. The implementation of the system in the context of Vortal e-marketplace is expected to be carried out in the near future.

Behind the creation of CIsea were wider objectives related to the identification of the critical factors contributing to company and project success. The indicators included in CIsea cover three main areas: company financial performance, operation performance, and quality of the proposals. CIsea generates outcomes based on individual indicators, and also determines an aggregate measure of overall performance. This aggregate measure is calculated using a weighted average of the indicators, where the weights attached to each indicator are chosen by the decision maker. This simplified weighting method to evaluate performance was intended to foster its acceptance among CI company managers. The use of more sophisticated quantitative methods to measure performance, such as Data Envelopment Analysis or multicriteria techniques, is in the agenda for future developments after implementing the system.

The remainder of this chapter is organized as follows. Section 4.2 focus on the selection of the key performance indicators to be included in CIsea. Section 4.3 presents the system developed, explaining how it interacts with e-marketplaces, its main features, and the type of performance results provided. The last section concludes and suggests directions for future research.

4.2 The selection of key performance indicators

Both GCs and SCs play an important role in the success of construction projects (Arditi and Chotibhongs, 2005; Palaneeswaran and Kumaraswamy, 2001), so it is crucial to establish a set of criteria through which the capabilities of both are correctly measured and judged. Therefore, the study concerning the key performance indicators (KPIs) to be used in the CIsea system represented a particularly important task. The methodology used to identify the most critical indicators consisted in two main steps, as explained next.

In the first step, we reviewed the literature focusing on the studies that discuss with particular detail the most appropriate criteria for the evaluation of GCs and SCs in order to support the decision making during the bidding process.

Concerning the prequalification criteria used to evaluate GCs, the study developed by Bubshait and Al-Gobali (1996) identified the most common prequalification criteria used in Saudi Arabia: contractor experience, financial stability, past performance, quality performance, project management capabilities, failed contracts, management staff availability, and contractor capacity. Ng and Skitmore (1999) concluded that in the United Kingdom the ten most common criteria to prequalify GC include: performance, fraudulent activity, progress of work, financial stability, competitiveness, management capability, failed contracts, company stability, relationship with client, standard of quality. El-Sawalhi et al. (2007), based on an extensive literature review, identified the most common criteria to evaluate contractors: financial stability, technical ability, experience, performance in terms of cost, time and quality, resources utilization, quality management, safety and environmental concerns. Concerning the criteria used to evaluate SC, only the study developed by Hartmann et al. (2009) examined the use of four criteria in the SC selection process in Singapore: price, technical know-how, quality, and cooperation. The study revealed that price is the criterion most often used.

From the literature review, we defined a list including the indicators considered more relevant concerning two categories of company attributes: financial performance indicators and operation performance indicators.

In the second step, we validated the indicators originally selected to be included in the CIsea platform, and completed them with other relevant indicators. This was accomplished by undertaking a meeting with Vortal members and practitioners from two companies, which are Vortal clients, with a deep knowledge of the Portuguese CI. As a result of this meeting, it was suggested to include another category of indicators in the CIsea system, related to the bid/tender attributes. This category is of particular interest to give insights concerning the competitiveness of the proposals presented, and to provide guidance to improve their quality. This is relevant for GCs in order to prepare accurate tenders to minimize the risk of unpredicted aspects during the job execution, and also for SC in order to increase the number of bids contracted. In addition, it was suggested to include an extra indicator concerning the accomplishment of the contractual conditions (called "Contract compliance" hereafter) in the list of operation indicators for both GC and SC evaluation. This dimension is perceived as essential to ensure the success of a project. It is generally believed that the lack of commitment with the contractual conditions is one of the critical weaknesses in the Portuguese CI, often leading to time deviations and additional costs.

It was made an effort to define a final set of indicators aligned with the specificities of the CI in Portugal, requiring information that could be easily collected by construction companies. We also tried to keep the indicators with a wide scope to allow future extensions for the evaluation of other agents involved in the CI value chain (e.g. consultants, owners, ultimate clients).

The final set of KPIs selected includes 13 indicators for SC evaluation and 10 for GC evaluation, divided into three main categories: financial performance, operation performance, and bid/tender attributes. Figure 4.1 and Figure 4.2 show the indicators considered within each major category for SC evaluation and GC evaluation, respectively.

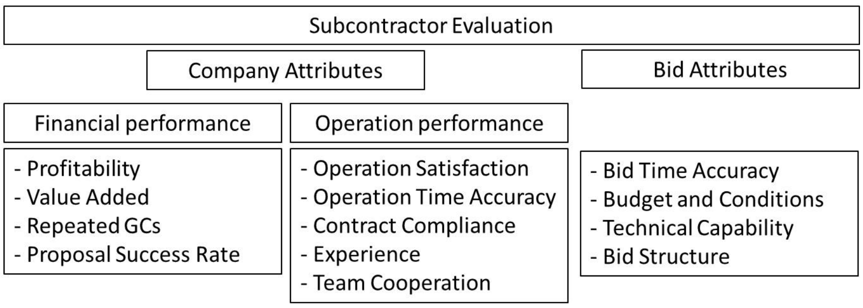

Figure 4.1: Indicators for Subcontractor evaluation

Concerning the financial performance, the indicators included to evaluate SC and GC were as follows: i) Profitability - to measure the profit of the company before tax and interest; ii) Value Added - to measure the contribution to Gross Value Added made by an individual employee. For SC evaluation, we include two additional indicators related to company reliability to complement the financial performance category: i) Repeated GCs - to measure the percentage of GCs repeated by a SC in the e-marketplace;

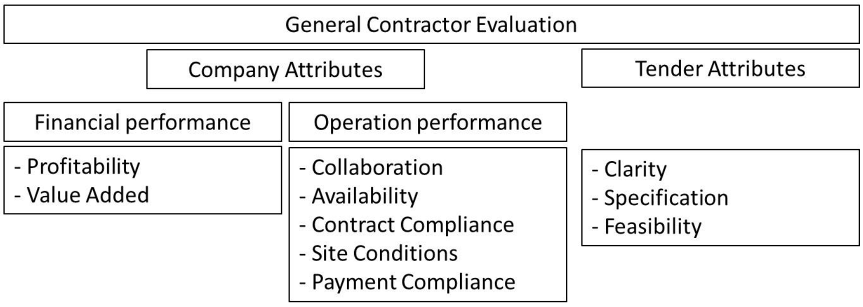

Figure 4.2: Indicators for General Contractor evaluation

ii) Proposal Success Rate - to measure the percentage of bids awarded in the e-marketplace. The financial performance indicators are quantitative indicators, measured on a continuous scale. Table 4.1 presents the formulas used to calculate such indicators.

Table 4.1: Financial indicators and respective formula

Indicators	**Formulas**	
Profitability	$PROF = \frac{P}{V} \times 100$	P-Profit before taxes (€) V-Value of sales (€)
Value Added	$VA = \frac{GVA}{N}$	GVA-Gross value added (10^3€) N-Total no. of employees
Repeated GCs	$RC = \frac{NRC}{NC} \times 100$	NRC-No. of repeated GCs NC-No. of GCs
Proposal Success Rate	$PSR = \frac{PA}{PP} \times 100$	PA-No. of proposals awarded PP-No. of proposals presented

Concerning operation performance, the indicators included to evaluate SC were as follows: i) Operation Satisfaction - to measure the satisfaction with the operation execution; ii) Operation Time Accuracy - to measure the satisfaction with the completion of operation within time; iii) Contract Compliance - to measure the satisfaction with the accomplishment of contractual conditions in terms of cost, time, human resources, materials, and equipments; iv) Experience - to measure the satisfaction with the know-how demonstrated by the employees; v) Team Cooperation - to measure the satisfaction with the cooperation between employees from the SC company and the other teams in the work site.

To evaluate operation performance of GC, five indicators were specified: i) Collaboration - to measure the satisfaction with the GC involvement during the operation; ii) Availability - to measure the satisfaction with the GC availability in terms of discussing unexpected situations, answering questions, participating in meetings, or finding consensual solutions; iii) Contract Compliance - to measure the satisfaction with the accomplishment of contractual conditions in terms of time, reworks, materials, and equipments; iv) Site Conditions - to measure the satisfaction with the conditions of the work site in terms of organization during the operation execution; v) Payment Compliance - to measure the satisfaction with the accomplishment of payment procedures in terms of liquidity terms, time taken to analyze invoices, and to deal with administrative procedures.

The third category corresponds to the proposals (bids/tenders) attributes. To evaluate bid proposals four indicators were specified: i) Bid Time Accuracy - to measure the satisfaction with the submission of the bid within time; ii) Budget and conditions - to measure the satisfaction with the budget and conditions proposed for the payment process; iii) Technical capability - to measure the satisfaction with the adequacy of the human resources, materials, and equipments; iv) Bid structure - to measure the satisfaction with the presentation and detail of the bid in terms of draws, equipments, and resources. For the evaluation of tenders were considered three indicators: i) Clarity - to measure the satisfaction with the tender comprehensibility in terms of works, due dates, and resources; ii) Specification - to measure the satisfaction with the presentation and detail of the tender; iii) Feasibility - to measure the satisfaction with the feasibility of schedules, prices, human resources, and works required.

The indicators related to operation performance and bid/tender attributes are qualitative variables. These indicators are measured using a Likert scale of discrete numbers from 1 to 10, where 10 means totally satisfied, 5-6 neither satisfied nor dissatisfied and 1 totally dissatisfied.

4.3 The performance assessment system (CIsea)

4.3.1 The evaluation process

The CIsea system was designed to be integrated within an e-marketplace. Figure 4.3 shows the main steps of the bidding process in an e-marketplace and the interaction between CIsea and the e-marketplace.

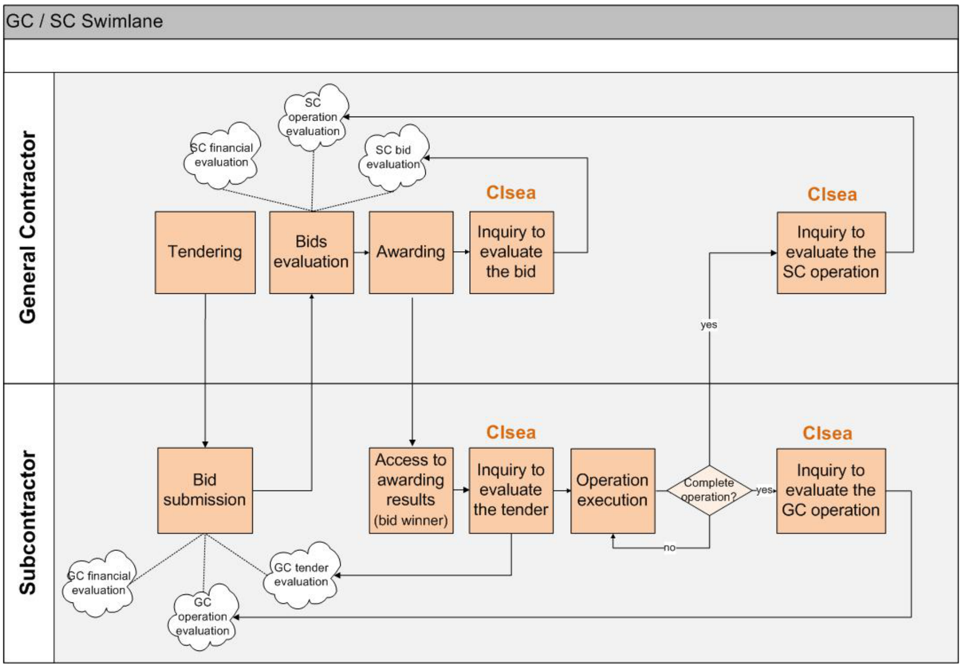

Figure 4.3: Interaction between CIsea and e-marketplace

The bidding process starts when the GC introduces a tender in the e-marketplace. All SCs registered in the e-marketplace are notified about the tender, and may submit a bid. At this stage, they can visualize information concerning the GC financial, operation and tender performance evaluations from previous jobs. Then, the GC analyzes all bids submitted in the e-marketplace in order to select the most appropriate. This is supported by the visualization of results concerning the SC financial, operation and bid performance evaluations from previous jobs, as explained in detail in the next section. After selecting the best bid and awarding the contract, the GC fulfills the questionnaire concerning the attributes of the bid selected. In turn, the SC contracted fulfills the questionnaire related to the tender attributes. The items included in the questionnaires correspond to the indicators related to the bid and tender attributes, as presented in Figures 4.1 and 4.2 for SC and GC, respectively. After job completion, the GC fulfills the questionnaire related to SC operation performance. In turn, the SC answers the questionnaire related to GC operation performance. The items included in the questionnaires correspond to the indicators related to the operation performance, as presented in Figures 4.1 and 4.2 for SC and GC, respectively. This procedure is repeated for each bid contracted through the e-marketplace.

Note that the data needed to calculate the indicators related to financial performance can be directly gathered from the e-marketplace. In particular, the data needed to determine Profitability and Value added is mandatory for SCs and GCs registered in the Vortal e-marketplace. The data needed to calculate Repeated GCs and Proposal Success Rate indicators is directly collected from the e-marketplace records.

4.3.2 The visualization of performance results

The CIsea plays a fundamental role to support the GC in the selection of the SC to be contracted among competitive bids. At this stage, the GC may visualize the performance results of the SCs that submitted a bid in terms of financial, operation and bid performance. The CIsea platform exhibits two types of performance results: i) the past results obtained in the jobs carried out between the SC and the GC; ii) the past results obtained in the jobs carried out between the SC and other GCs also registered in the e-marketplace. CIsea is also a powerful tool for the SC as it enables to visualize the performance results of the GCs that submitted a tender in terms of financial, operation and tender performance. The same type of results are presented to SC: i) the past results obtained in the jobs carried out between the GC and the SC; ii) the past results obtained in the jobs carried out between the GC and other SCs also registered in the e-marketplace.

The CIsea system enables to have a view of how a company is performing in several indicators. However, the large variety of indicators available makes it difficult to gain an overall view of performance, as it is possible to perform very well on some indicators but poorly on others. Hence, a measure that gives an idea of the overall performance when several indicators are considered is also generated by CIsea. The aggregate measure reflects the evaluations given by other companies to a particular company. Therefore, it only aggregates the indicators related to operation and bid/tender attributes.

The aggregate measure of overall performance (OP) is calculated based on a weighted average, taking into account the decision maker preferences. The weights assigned to each individual indicator range between 0% and 100%, summing 100% for the total set of indicators. In particular, weights equal to zero exclude the respective indicators from the assessment. Note that by default, the CIsea platform calculates the aggregate measure based on a simple arithmetic average with all indicators with equal weight. This can be customized by the decision makers, that can assign different weights to reflect their preferences.

The CIsea system also enables the visualization of performance results at three different levels. The first level corresponds to a visualization with greater detail at the job level, which includes the operation and bid/tender perspectives. For the job level, the result of each indicator is obtained from the questionnaires fulfilled by GCs and SCs. The second level corresponds to visualize results by projects, which corresponds to aggregate the evaluations of all jobs belonging to the same project. The third level corresponds to visualize results by companies, which corresponds to aggregate the evaluations of all projects related to the same company. The company level includes additionally the visualization of financial indicators.

To show how the performance results are presented in CIsea, Figure 4.4 illustrates the performance evaluation that can be visualized by a GC in relation to a particular SC at the company level. The results for each individual indicator related to operation and bid attributes appear in the form of horizontal bars. CIsea exhibits two distinct bars for each individual indicator. The green bar displays the arithmetic mean of all evaluations given to the SC by the GC that is visualizing these results, named "company to company" (C2C) evaluation in the platform. The orange bar displays the arithmetic mean of all evaluations given to the SC by registered GCs, named "market to company" (M2C) evaluation. The bar charts are useful to quickly grasp the variability among the indicators as well as their interval limits (e.g. maximum and the minimum values). The results of the financial indicators are also reported in the first panel of Figure 4.4. Figure 4.4 also exhibits the aggregate measures of overall performance, and the sample size details on the bottom.

Concerning the evaluations given by the GC to this particular SC, it can be observed that the overall performance score is equal to 8.7. This score is obtained from a sample corresponding to all jobs evaluated by the GC (40 in this case). In terms of the individual indicators, we observe that the worst

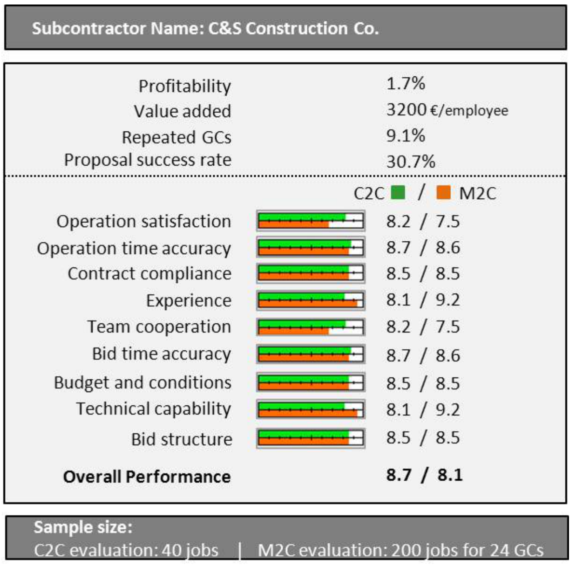

Figure 4.4: Performance results of a Subcontractor available in CIsea

performance occurs for Experience and Technical Capability indicators, with an average result of 8.1. The best performance occurs for Operation Time Accuracy and Bid Time Accuracy indicators, with an average result of 8.7.

Concerning the evaluations given by all registered GCs to this particular SC, it can be observed that the overall performance score is equal to 8.1. This score is obtained from a sample corresponding to 200 jobs, evaluated by 24 distinct GCs. In terms of the individual indicators, we observe that the worst performance occurs for Operation Satisfaction and Team Cooperation indicators, with an average result of 7.5. The best performance occurs for Experience and Technical Capability indicators, with an average result of 9.2.

Additionally, CIsea has the ability to report historical performance results for a given SC or GC. This functionality is very useful to analyze trends in results over time and to diagnose areas in need of improvement for a given company. The historical results are presented in the form of vertical bars and can reflect the results of individual indicators or overall measures of performance. The CIsea system enables to visualize results by different time windows, i.e. annual, semester, quarterly and monthly results. A company can see its historical performance results, or the historical results of the other companies involved in the bidding process.

To illustrate how historical performance results are presented in CIsea, Figure 4.5 shows the overall scores that can be visualized by a GC in relation to a particular SC, concerning the period between 2006 and 2011. In particular, the green bars correspond to the overall scores related to the evaluations of the SC given by the GC (C2C evaluation). The orange bars are the overall scores related to the evaluations of the SC given by all registered GCs (M2C evaluation). The details button shown in Figure 4.5 provides the details related to the sample size that underlie the calculation of the overall performance score in each bar.

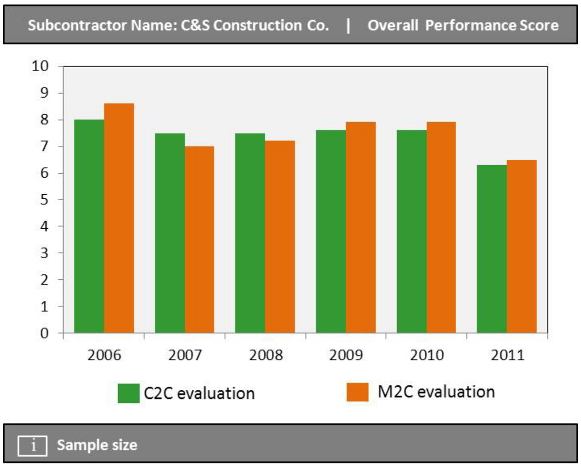

Figure 4.5: Historical results by overall performance scores

All the graphs that report the performance results aim to make available to company managers a powerful tool to support decision making in e-marketplaces, as well as to guide sustainable improvement strategies.

4.4 Summary and conclusions

Selecting "the right person for the right job" is crucial to ensure the success of the construction project (Palaneeswaran and Kumaraswamy, 2000). CIsea aims to support the selection of the most suitable company for a given work based on a set of diverse and relevant key performance indicators. The final set of indicators includes an innovative perspective, not previously discussed in the literature, related to bid/tender attributes, which is essential to monitor the quality and the competitiveness of the proposals. One of the most important features of CIsea platform is to be able to accommodate bilateral evaluations between companies. This enables a better understanding of the GC-SC relationship, and consequently to identify possible causes for the success/failure of a construction project. Another key aspect that distinguishes CIsea from the existing platforms is to allow the incorporation of individual decision maker preferences in terms of the relative importance of the criteria used for performance assessments when estimating an overall performance score.

For future developments, the first step corresponds to implement CIsea in the context of Vortal e-marketplaces. After implementing CIsea, there are other topics for future research. Although CIsea is designed to support the bidding process between GCs and SCs, the system developed can be easily adapted to help the bidding between owners and GCs. This could contribute to a wider assessment of the entire construction project, which is particularly important to guide companies towards more productive levels, and to improve the performance of the CI as a whole. Another major development is to incorporate more sophisticated methods for performance measurement in the assessments available in the CIsea platform. In particular, multicriteria techniques such as the Analytical Hierarchy Process to assist the decision-maker to introduce the relative importance (weighting) of each indicator, according to its relevance.

Finally, it would be particularly interesting to undertake benchmarking exercises using the CIsea platform. This would allow comparing the performance of a specific company with other companies on several indicators. To allow contextualized comparisons, it would be interesting to enable choosing the sample for comparison, such as companies with the same headquarter location, the same type of ownership (public or private), or the full sample of competitors. In this context, the Data Envelopment Analysis technique could be used to compute an overall performance measure, taking into account the performance of other companies in the sample.

Chapter 5

Performance assessment of construction companies integrating Key Performance Indicators and Data Envelopment Analysis

5.1 Introduction

The increasing competitiveness of Construction Industry (CI) motivated companies to assess performance and implement efficiency improvement strategies in order to obtain competitive advantage. Benchmarking has become a common practice in the sector. Benchmarking was first introduced by Camp (1989) who defined it as “the continuous process of measuring products, services, and practices against the toughest competitors or those companies recognized as industry leaders”. The main purpose of benchmarking is to improve performance on a continuous basis by incorporating best practices in the business process. More recently, in several countries, benchmarking systems were specially developed for the CI. These systems, usually available in web platforms, typically analyze company performance based on a set of key performance indicators (KPIs), which consist on ratios representing key aspects of company activity.

In spite of the generalized acceptance of using a set of KPIs to evaluate company performance, there are a few theoretical and empirical limitations associated with their use. Each individual indicator examines only a part of the company activity, so a comprehensive performance evaluation must be based on the analysis of several indicators. Therefore, it may be difficult to gain an overall performance view, as the number of indicators that can be computed for each company may be unmanageably large. Even assuming that a subset of KPIs is identified, the ranking of companies is still impaired, particularly because it is unlikely that indicators associated to different dimensions indicate a similar level of performance for each company. In general, the multi dimensional problem is solved by the construction of a synthetic indicator, obtained by normalizing and averaging all scores assigned to an organization on the different criteria.

Another limitation of using a set of KPIs is that they cannot be used in a straightforward manner to establish improvement targets. This is because each single indicator has to be compared to some benchmark value, without regarding the remaining aspects of the company activity that are not accounted for in that indicator. Although any particularly poor value for an indicator identifies an aspect of the company activity requiring improvement, the target levels cannot be estimated with confidence, as achieving a target for one indicator may have implications on other dimensions of company activity.

The main contribution of this chapter is to develop a methodology for assessing company overall performance, that can complement the information provided by traditional CI assessments. The methodology proposed combines KPIs with DEA, and was designed with the purpose to be useful to all organizations involved in benchmarking routines. One of the advantages of the DEA technique is to allow aggregating multiple dimensions of company activity, evaluated by several KPIs, into a single summary measure of performance. This measure is obtained in such a way that shows each DMU in the best possible light. This enables comparing companies in terms of relative performance. Moreover, DEA identifies a subset of efficient organizations, considered as examples of best practice, and specifies improvement targets for inefficient companies, which take into account all KPIs considered.

To illustrate how DEA can be integrated with the results of benchmarking platforms, consisting on benchmark scores for each KPI, the Portuguese benchmarking platform for the CI, icBench, was used. The assessment of a sample of twenty Portuguese contractors described in this study adopted two different perspectives: organizational performance and operations performance. To assist strategy definition in a more realistic way, two types of DEA models were used in this analysis: one allows weights to vary freely, and the other includes weight restrictions to accommodate managers' preferences. Furthermore, to enable suggesting targets for all organizations, manager opinion was used to specify virtual units, which were included in the efficiency assessment to define a practical frontier located beyond the performance levels of the original DEA frontier.

The remainder of this chapter is organized as follows. Section 5.2 describes the methodology used in this chapter. Section 5.3 describes the data used and the value judgements introduced in the performance assessment. Section 5.4 discusses the results obtained. The last section presents the conclusions and recommendations for future research.

5.2 Methodology

5.2.1 The use of DEA to construct composite indicators

A composite indicator is a mathematical aggregation of a set of sub-indicators for measuring multi-dimensional concepts that cannot be captured by a single indicator (OECD, 2008). Composite indicators have increasingly been accepted as useful tools for performance comparisons, benchmarking, policy analysis and public communication in various fields such as economy, environment and society (OECD, 2008). There are various methods used to aggregate the indicators in a composite indicator, such as DEA or multicriteria decision analysis. The use of DEA to estimate composite indicators has gained popularity in recent years due to its ability to aggregate multiple dimensions of company activity, evaluated by several indicators, into a single summary measure of performance. This measure is obtained through linear programming, and shows each company in the most favorable light. The main feature of the DEA assessment is that each company can choose its own weighting system in order to reflect its strengths and strategic preferences.

The research line was initiated by Lovell et al. (1995) that proposed a DEA model to evaluate countries' performance based on four indicators related to services provided to citizens. In the DEA model, it was used a unitary input for each country, which can be interpreted as a "helmsman" that pursues the four service indicators. Afterwards, Cherchye et al. (2004) popularized the use of DEA for the estimation of composite indicators, by proposing the use of a simplified formation named "benefit of the doubt" weighting. The main difference between a traditional DEA analysis and the construction of a composite indicator, as proposed by Cherchye et al. (2004), is that the latter only looks at achievements, without explicitly taking into account the resources used. The rationale of using a DEA model to obtain a composite indicator is to aggregate a set of KPIs into a single summary measure of performance. The linear programming model for deriving the composite indicator for a DMU j_o proposed by Cherchye et al. (2004) is shown in (5.1).

$$\begin{aligned}
&c_{j_o} = \max \sum_{r=1}^{s} u_r \; y_{rj_o} \\
&\text{subject to} \\
&\sum_{r=1}^{s} u_r \; y_{rj} \leq 1, \;\; j = 1, \ldots, n \\
&u_r \geq \epsilon, \;\; r = 1, \ldots, s
\end{aligned} \tag{5.1}$$

In model (5.1), y_{rj} corresponds to the value of the output indicator r ($r = 1, \ldots, s$) in company j ($j = 1, \ldots, n$), with higher values corresponding to better performance. As stated by Cherchye et al. (2004), model (5.1) is equivalent to the constant returns to scale DEA input oriented model (2.8) with all indicators considered as outputs and a "dummy input" equal to one for all DMUs. The weights u_r ($r = 1, \ldots, s$) are the variables of model (5.1). Model (5.1) determines the weights u_r that give the highest possible score for each unit assessed, keeping the scores of all other units less than or equal to one when evaluated with similar weights. The composite indicator score $c^*_{j_o}$ of DMU j_o is between 0 (worst) and 1 (benchmark).

The rationale for using model (5.1) to construct a composite indicator is easy to explain: since it is difficult to identify a priori a set of weights that all DMUs would agree that reflects adequately the relative importance of each indicator y_r, we let each DMU select its own weights, such that its composite indicator is as high as possible compared to the composite indicator of other DMUs evaluated with similar weights. If we impose an upper bound of one to the highest composite indicator obtained across all DMUs, a value of the composite indicator equal to one signals best performance (i.e., benchmarks). Under these conditions, if a DMU does not achieve the maximum score, even when evaluated with a set of weights that intends to maximize its performance score, it provides irrefutable evidence that other DMUs performed better during the period considered.

Although model (5.1) allows the specification of the indicators' weights recurring to optimization, it is possible to incorporate in the model expert opinion about the relative importance of individual indicators. This information can be incorporated in model (5.1) by imposing restrictions to indicator weights, as illustrated in section 5.2.3.

5.2.2 Practical frontier in DEA models

The original DEA model allows to identify targets for inefficient DMUs suggesting improvements to their efficiency. For efficient DMUs no further improvement can be indicated. Nevertheless, if the industry is to improve as a whole, it is important to identify targets even for efficient companies. The Practical DEA (P-DEA) model developed by Sowlati and Paradi (2004) enables to create new DMUs, more efficient than those defining the empirical DEA frontier, to form a new improved frontier. The new "practical frontier" can be established based on expert opinion. This requires specifying possible variations in the input and output levels for the efficient DMUs, as well as the maximum efficiency improvement (δ) considered feasible. With this information, it is possible to obtain for each of the DMUs originally classified as efficient, a new DMU (called virtual DMU), satisfying the restrictions above. In the case of an assessment involving only outputs, the outputs of the virtual DMU can be obtained from a simplified P-DEA model as presented in (5.2). The decision variables of this model are the outputs of the new DMU ($\tilde{y}_{rj_o}$) and the output weights (u_r).

$$
\begin{aligned}
&\hat{c}_{j_o} = \max \sum_{r=1}^{s} u_r \ \tilde{y}_{rj_o} \\
&\text{subject to} \\
&\sum_{r=1}^{s} u_r \ y_{rj} \leq 1, \ j = 1, \ldots, n \\
&1 \leq \sum_{r=1}^{s} u_r \ \tilde{y}_{rj_o} \leq 1 + \delta, \ j = 1, \ldots, n \\
&L_{rj_o} \leq \tilde{y}_{rj_o} \leq U_{rj_o}, \ r = 1, \ldots, s \\
&u_r \geq \varepsilon, \ r = 1, \ldots, s
\end{aligned} \tag{5.2}
$$

The objective function is to maximize the efficiency of the virtual DMU, subject to the additional restrictions imposing that the outputs of the virtual DMU are within the limits defined by the decision maker (L_{rj_o} and U_{rj_o}), and that the virtual DMU has an efficiency score greater or equal to one, with an upper bound equal to $1 + \delta$.

Model (5.2) is non-linear, but multiplying the restriction $L_{rj_o} \leq \tilde{y}_{rj_o} \leq U_{rj_o}$ by the weight u_r, and making the variable substitution $u_r \ \tilde{y}_{rj_o} = p_{rj_o}$ in (5.2), it becomes a linear programming model. The resulting model (5.3) should be run for each efficient DMU.

$$
\begin{aligned}
&\hat{c}_{j_o} = \max \sum_{r=1}^{s} p_{rj_o} \\
&\text{subject to} \\
&\sum_{r=1}^{s} u_r \ y_{rj} \leq 1, \ j = 1, \ldots, n \\
&1 \leq \sum_{r=1}^{s} p_{rj_o} \leq 1 + \delta, \ j = 1, \ldots, n \\
&u_r \ L_{rj_o} \leq p_{rj_o} \leq u_r \ U_{rj_o}, \ r = 1, \ldots, s \\
&u_r \geq \varepsilon, \ r = 1, \ldots, s
\end{aligned} \tag{5.3}
$$

After obtaining the solution to this linear programming model, the outputs of the virtual DMU can be retrieved using the expression $\tilde{y}_{rj_o} = p^*_{rj_o}/u^*_r$. The symbol $(*)$ corresponds to the values of the variables at the optimal solution. Note that using the P-DEA model (5.3), for some of the DMUs originally considered efficient could happen that no efficiency improvements are possible. After adding the virtual DMUs to the original sample, the DEA analysis can be done with reference to a new practical frontier. In essence, this procedure allows moving the empirical frontier constructed only with the original DMUs towards more productive levels.

5.2.3 Illustrative example

The empirical assessment described in this chapter started with the evaluation of companies' performance using the composite indicator model (5.1). This was followed by another assessment also using model (5.1) but this time including weight restrictions to enable a more realistic performance assessment for each company, aligned with decision maker opinion. Then, a Practical DEA model was used to create virtual DMUs based on the efficient DMUs identified in the weight restricted composite indicator model. Finally, the composite indicator model including weights restrictions was run again including the original companies and the virtual companies created in order to provide directions for improvement to all companies, even for the efficient ones.

To illustrate the methodology applied we use an example consisting of five DMUs. These are assessed considering two outputs (Y_1) and (Y_2) that represent performance indicators. Suppose that, in the context of CI companies, (Y_1) represents the number of winning bids per employee and (Y_2) the profit, in euros, per employee. Table 5.1 lists the values of the outputs for each DMU, and Figure 5.1 illustrates the production possibility set.

Table 5.1: Outputs of the DMUs considered in the illustrative example

DMUs	No. of winning bids per employee (Y_1)	Profit per employee (Y_2)
A	0.18	85
B	0.125	125
C	0.10	80
D	0.13	60
E	0.025	65

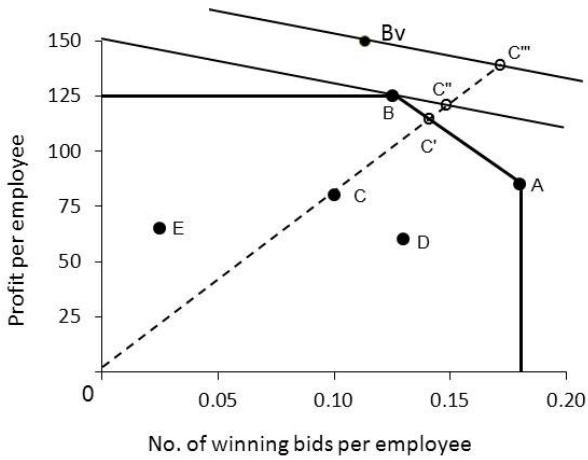

Figure 5.1: Representation of the production possibility set including a virtual unit

From Figure 5.1, it is possible to visualize that units A and B define the efficient frontier that envelops all other units. Therefore, they both achieve a composite indicator, estimated using model (5.1), equal to 100%. The remaining units C, D, E are inefficient. In particular, the composite indicator of unit C is 70.7%, which is graphically represented by the ratio OC/OC'. With complete weight flexibility, the ratio between the weights of output Y_1 and output Y_2 for unit C obtained at the optimal solution to model (5.1) is equal to $u_{Y1}/u_{Y2} = 732.6$. This trade-off can be interpreted that winning one bid per employee is perceived as equivalent to obtaining a profit per employee of €732.6. However, managers may express a preference towards obtaining higher profitability levels instead of winning bids. In quantitative terms, this could be transformed into a weight restriction imposing that for instance $u_{Y1}/u_{Y2} \leq 200$, meaning that winning one bid per employee cannot be valued more than a profit of €200 per employee. With this restriction, DMU A is no longer efficient, as its profitability level is quite small (i.e., it is €40 below the value observed in B, and this decrease in profit is not compensated by the increase in the number of winning bids per employee). Therefore, in an assessment with the weight restriction described above, only unit B would be considered efficient. The efficiency evaluation of unit C with this weight restriction would result in a comparison with a "value frontier" represented in Figure 5.1 by the segment tangent to the efficient frontier in B. DMU B is now the peer for all DMUs analyzed. The efficiency score of unit C decreased to 66.7%, which is represented by the ratio OC/OC''.

In order to create a new practical frontier, the first step consists on creating a virtual DMU based on the DMU initially considered as efficient (B) using model (5.3). It was considered that the efficiency of the virtual DMU (Bv) could increase 4% ($\delta = 0.04$) in relation to the original frontier and its outputs could decrease no more than 10% ($L_{rBv} = 0.9 \times y_{rB}$) and increase no more than 20% of the original output values ($U_{rBv} = 1.2 \times y_{rB}$). At the optimal solution to model (5.3), we would obtain the output levels of the virtual DMU (Bv) corresponding to $Y_1 = 0.113$ and $Y_2 = 150$.

In the new DEA assessment with the original sample and the virtual unit, Bv would be the only unit considered to be full efficient with the weight restrictions imposed, so it would become the peer for all units analyzed. As the standard of excellence became higher with the specification of the virtual frontier, the efficiency score of unit C would decrease to 57.9%, which is represented by the ratio OC/OC'''.

5.3 The data

5.3.1 Organizational performance perspective

The data used in this study came from the icBench platform, and corresponds to the year of 2005 (for further details see Costa and Horta, 2007). In particular, a sample of 50 Portuguese contractors was invited to participate in the icBench initiative. The companies were selected for their awareness of Information Technologies and use of advanced management methods. Twenty contractors voluntary accepted the invitation to participate. The study related to the organizational performance measurement model used the data collected from this set of twenty contractors. This sample includes some of the Portuguese leading construction companies in terms of value of sales and number of employees, representing 19.4% of the total sales and 2.5% of the total employment of the CI. The majority of them are located in the North of Portugal (65%) and the remaining are from the Center and the South of the country.

The organizational performance model was designed to provide an overall picture of the organizational performance of a contractor. This model included relevant indicators mainly reflecting financial aspects: i) Productivity (PROD) - to determine the value-added per employee; ii) Profitability (PROF) - to measure the profit of the company before tax and interest; iii) Hanging Invoice (HI) - to determine the total invoice not received; iv) Accident Frequency Rate (AFR) - to measure the number of reportable accidents per 10^6 worked hours; v) Sales Growth (SG) - to determine the evolution of market share. The items included in the questionnaires of the icBench platform, from which the data for the organizational performance model was collected, are presented in Table 5.2. Table 5.3 presents the formulas to obtain the indicators used.

Table 5.2: Indicators of the organizational performance model and the corresponding items included in the questionnaire

Organizational Performance Indicators	**Items Inquired**
Productivity	Value of sales
	Average number of full-time equivalent employees
	Value of services and suppliers
	Value of raw materials
Profitability	Profit before taxes
Hanging Invoice	Accounts receivable
Accident Frequency Rate	Number of reportable accidents
	Number of person-hours worked including own and subcontracted employees
Sales Growth	Value of sales in the previous year

Table 5.3: Indicators of organizational performance

Indicators	**Formula**	
Productivity	$PROD = \frac{(V_S - V_M - V_{SSE})}{N_E}$	V_S-Value of sales (€) V_M-Value of raw materials(€)
Profitability	$PROF = \frac{P}{V_S} \times 100$	V_{SSE}-Value of services and supplies (€) N_E- Average no. of full-time employees
Hanging Invoice	$HI = \frac{I}{V_S} \times 100$	P-Profit before taxes (€) I-Accounts receivable (€)
Accidents Frequency Rate	$AFR = \frac{N_{ac}}{N_{ph}} \times 100$	N_{ac}-Number of reportable accidents N_{ph}-Number of person-hours worked including own and subcontracted employees
Sales Growth	$SG = \frac{(V_S - V_{S-1})}{V_{S-1}} \times 100$	V_{S-1}-Value of sales in previous year (€)

The descriptive statistics of the variables are presented in Table 5.4. In this model, the DMUs refer to contractors. It can be concluded that the set of companies is quite diverse, which can be seen from the relatively high values of standard deviation.

In benchmarking platforms the original KPI values are converted to the corresponding benchmark scores. Figure 5.2 illustrates this procedure using

Table 5.4: Descriptive statistics of original output data for organizational model

Output Variables	Mean	Standard Deviation	Minimum	Maximum
PROD	35787	16900	7804	76689
PROF	1.6	4.4	-10.5	7
HI	34.4	20.9	3.9	93.8
AFR	125.6	326.7	2.9	1 467
SG	17.5	29	-32	71.4

icBench data for the five indicators used. It shows the conversion of the original KPI measures (vertical axis) to the benchmark scores (horizontal axis) for the twenty companies. The benchmark data is measured in percentage, where 0% corresponds to the worst performance achieved and 100% corresponds to the best level of performance. Note that for the Hanging Invoice (HI) and Accident Frequency Rate (AFR) indicators, a higher value means worst performance. The conversion to benchmark values was done in such a way that a higher benchmark score corresponds to better performance. Therefore, the curve of these indicators corresponds to a descending line with negative slope, whereas for the other indicators the slope is positive.

To enable the integration of the results of benchmarking platforms with the DEA, we used a composite indicator model where the outputs of the model are the benchmark scores for the five indicators reported on Table 5.3.

Since four of the twenty contractors did not provide data concerning Sales Growth (SG), the data collected had four missing values for this output. These values needed to be dealt with since it was not desirable to exclude any company from the evaluation. The missing data of the output variable was replaced by zero, following the procedure proposed by Kuosmanen (2002). As explained by Portela and Camanho (2007) modeling missing data in this way corresponds to assume that a DMU cannot weight the factors that are missing. Note that if the value of a certain variable is missing, we do

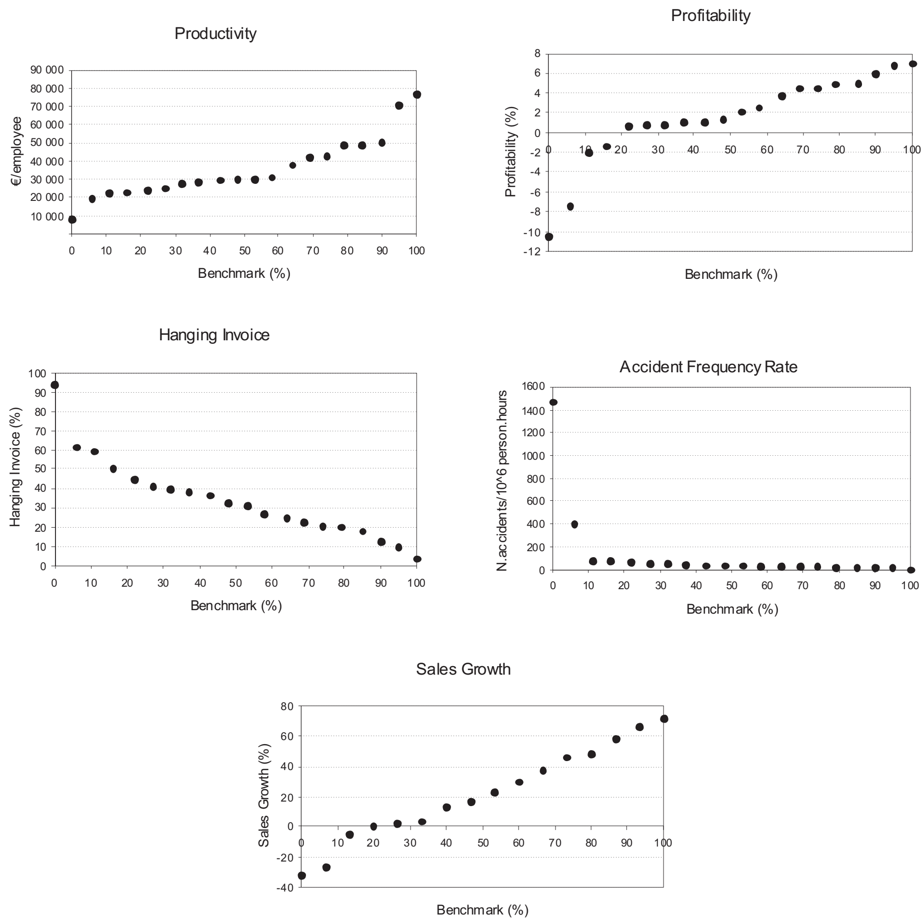

Figure 5.2: Conversion of the original output data to the corresponding benchmark scores

not know if the company performed well or not on that aspect. If the missing variable corresponds to an area where the company is underperforming, then in an assessment without missing data the unit would choose not to weight the underperforming factor. Therefore, the efficiency score obtained following this procedure would be identical to the score obtained in an assessment with all data. On the other hand, if the missing variable corresponds to an aspect with good performance, then in an assessment without missing data the unit could choose to weight that factor and its efficiency could improve in relation to the assessment with missing data. Therefore, the assessment of a unit with missing values always implies a lower (or equal) efficiency score that the assessment with complete data.

5.3.2 Operations performance perspective

The sample used in the operations performance measurement model only represented a subset of eight companies from the original twenty contractors, since the others did not provide sufficient operations data. The entire sample consisted of sixty operations, although the number of operations associated to each company varied.

The operations performance model has the objective of evaluating the operations finished by each contractor. It focuses primarily on the quality of the relationships between the contractor and the clients, construction materials suppliers, subcontractors and inspection teams. These aspects are very important due to their influence on the success or failure of a project. The operations model also considered a measure of cost deviation, since it represents a critical aspect in this sector. Although it would be advantageous to consider in the assessment other measures such as time or quality of the construction projects, the data provided by companies was insufficient to include them in this analysis.

The operations performance model considered the following indicators: i) Contractor satisfaction with customer cooperation (CS-CC) - to determine the contractor overall satisfaction level with customer involvement; ii) Contractor satisfaction with payments availability (CS-PA) - to determine the contractor overall satisfaction level with customer accomplishment of the predicted payment terms; iii) Contractor satisfaction with cooperative work (CS-CW) - to determine the contractor overall satisfaction level regarding the teams involved in the project; iv) Cost predictability (CP) - to measure the reliability of the construction cost estimates. The items included in the questionnaires of the icBench platform, from which the data for the operations performance model was collected, are presented in Table 5.5.

Table 5.5: Indicators of the operations performance model and the corresponding items included in the questionnaire

Operations Performance Indicators	**Items Inquired**
Contractor Satisfaction with Customer Cooperation	Quality of the project and contractual conditions
	Availability to discuss unexpected situations
	Availability to discuss change to orders
	Time taken to respond to questions
	Active participation in the process
	Overall satisfaction
Contractor Satisfaction with Payments Availability	Satisfaction with billing procedures
	Time taken for the analysis of invoices
	Time taken to deal with administrative procedures
	Accomplishment of the deadlines
	Overall satisfaction
Contractor Satisfaction with Cooperative work	Adequate professional competence
	Availability to participate in meetings
	Availability to find consensual solutions
	Time taken to respond to questions
	Involvement in cooperative work
	Overall satisfaction
Cost Predictability	Value of the proposal including errors and omissions
	Value of the final proposal including reworks

Note that indicators related to contractor satisfaction are measured using a Likert scale of discrete numbers from 1 to 10, whereas cost predictability is measured on a continuous scale. Table 5.6 presents the formulas used to obtain the indicators considered in the model.

The descriptive statistics for the indicators described in Table 5.6 are presented in Table 5.7. The indicator with greater variability is cost predictability (CP).

Table 5.6: Indicators of operations performance

Indicators	Formula	
	The score is the rating from the contractor survey, completed after finishing an operation, using a Likert scale (from 1 to 10), where:	
CS-CC		10 - Totally satisfied
CS-PA		5/6 - Neither satisfied nor dissatisfied
CS-CW		1 - Totally dissatisfied
CP (%)	$CP = \frac{Vf_r - Vi_{eo}}{Vi_{eo}} \times 100$	Vf_r- Value of the final proposal with reworks (€)
		$Vi_{eo}-$ Value of the proposal with errors and omissions (€)

Table 5.7: Descriptive statistics of original output data for contractors' operations model

Output variables	Mean	Standard Deviation	Minimum	Maximum
CS-CC	6.9	1.5	1	9
CS-PA	6.6	2.7	0[1]	10
CS-CW	7.3	1.5	3	10
CP	14.2	39.6	-100	203.9

[1]missing value

Following a similar procedure to the one used in the organizational performance assessment, the outputs used in the operational composite indicator model correspond to the benchmark scores of each indicator. The DMUs refer to the operations undertaken by the contractors. The only indicator with missing values is CS-PA since company C13 did not provide the data corresponding to its three operations. The replacement of missing values followed the procedure previously described.

5.3.3 Value judgements

The performance of all contractors was first evaluated using the composite indicator (5.1), and then using model (5.1) including weight restrictions. The weight restrictions were defined based on the value judgments expressed by construction experts.

For the organizational perspective, it was stated that all indicators should be considered in the evaluation, such that the overall efficiency measure should have at least 10% of the weight associated to each indicator. Furthermore, three indicators (productivity, profitability and hanging invoice) were perceived as being more important than the others, such that their total weight should be at least 70%.

For the operations perspective, the indicators concerning the quality of the relationships between the contractor and the other entities (i.e., satisfaction with customer cooperation, satisfaction with payments availability and satisfaction with cooperative work) were considered key aspects, such that their overall weight should be at least 80%. Similarly, it was considered that all indicators should be given a weight of at least 10%.

These value judgments expressed by experts where converted into weight restrictions in the form of assurance regions type I, as presented in section 2.2.4, to be added to the composite indicator model, as shown in Table 5.8. Note that since the outputs of the composite models correspond to normalized values, ranging from 0 to 100, the relative importance expressed as percentages could be directly converted to the bounds of the ratios between the weights. Note that in the organizational perspective, $r = 1, \ldots, 5$ correspond to the outputs productivity, profitability, hanging invoice, accidents frequency rate and sales growth, respectively. In the operations perspective, $r = 1, \ldots, 4$ correspond to satisfaction with customer cooperation, satisfaction with payments availability, satisfaction with cooperative work and cost predictability, respectively.

A P-DEA model was then used to create virtual DMUs based on the efficient DMUs identified in the weight restricted composite indicator model. The P-DEA model was defined based on the value judgments of construction experts concerning the allowable ranges for the change to the output levels

Table 5.8: Mathematical forms of the weight constraints used in the two perspectives

Organizational perspective	Operations perspective
$\frac{u_r}{\sum\limits_{r=1}^{5} u_r} \geq 0.10, r = 1, ..., 5$	$\frac{u_r}{\sum\limits_{r=1}^{4} u_r} \geq 0.10, r = 1, ..., 4$
$\frac{u_1+u_2+u_3}{\sum\limits_{r=1}^{5} u_r} \geq 0.70$	$\frac{u_1+u_2+u_3}{\sum\limits_{r=1}^{4} u_r} \geq 0.80$

and the potential increase in the efficiency of the best practice DMUs.

For the organizational perspective, it was stated that all the outputs (productivity, profitability, hanging invoice, accidents frequency rate and sales growth) may increase up to a maximum of 20% and decrease no more than 15% of the original value for each indicator. Furthermore, an increase in efficiency of 4% was perceived to be a realistic target.

For the operations perspective, it was suggested that all the outputs (contractor satisfaction with customer cooperation, contractor satisfaction with payments availability, contractor satisfaction with cooperative work and cost predictability) could increase up to a maximum of 30% and decrease no more than 10% of the original values. Similarly, a 4% increase in efficiency levels was considered realistic.

The restrictions corresponding to the integration of these value judgments in the P-DEA model (5.3) are presented in Table 5.9. Note that for both perspectives an additional constraint was specified stating that the outputs of the virtual DMUs cannot be greater than 100%, since all outputs are benchmark values.

Table 5.9: Parameters used in the P-DEA model in the two perspectives

Organizational perspective	Operations perspective
$\delta = 0.04$	$\delta = 0.04$
$(1 - 0.15)y_{rj_o} \leq \tilde{y}_{rj_o} \leq (1 + 0.20)y_{rj_o}, \forall r$	$(1 - 0.10)y_{rj_o} \leq \tilde{y}_{rj_o} \leq (1 + 0.30)y_{rj_o}, \forall r$
$\tilde{y}_{rj_o} \leq 100, \forall r$	$\tilde{y}_{rj_o} \leq 100, \forall r$

Finally, the composite indicator model (5.1) including the weights restrictions shown in Table 5.8 was run again including the twenty original companies and the virtual companies created. The software used to run the composite indicator models and the P-DEA model was the Advanced Integrated Multidimensional Modeling Software (AIMMS) from Paragon Decision Technology.

5.4 Results and discussion

5.4.1 Organizational performance assessment

The benchmark data and the KPIs average for each company are reported on Table 5.10. The performance results related to the organizational perspective are reported on Table 5.11. The first column corresponds to composite indicator model (5.1), the second and third columns correspond to composite indicator model (5.1) with the restrictions specified in Table 5.8. The second column only includes the original DMUs whilst the third column also includes the virtual DMUs. The symbol ($*$) indicates an assessment without virtual DMUs, and the symbol ($**$) an assessment including virtual DMUs.

To explore the face validity of both standard and restricted composite indicator results we explored in detail the KPI scores of the DEA efficient companies. With the composite indicator model, ten companies were considered efficient. Analyzing Tables 5.10 and 5.11, it is possible to observe that all DEA efficient companies achieved in at least one KPI dimension, one of the first five positions in the ranking. In terms of overall performance, measured by the KPIs average, all DEA efficient companies are in the top positions of the ranking (i.e., up to the 11th position). The Spearman rank-order correlation between standard composite indicator scores and KPIs average is equal to 0.860 (p-value equal to 0.000), which means that the DEA scores and the KPIs average are strongly associated. The difference between the

Table 5.10: Contractor benchmark data including virtual DMUs and respective KPIs average

DMU	PROD	PROF	HI	AFR	SG	KPIs average
C1	37	27	43	0	7	22.7
C2	64	58	48	100	100	74.0
C3	79	48	64	74	93	71.7
C4	22	53	22	11	80	37.6
C5	16	0	32	64	20	26.4
C6	27	37	74	32	27	39.3
C7	84	69	37	27	60	55.4
C8	48	32	79	37	73	53.9
C9	43	95	90	85	0	78.3
C10	58	16	58	53	33	43.7
C11	95	74	27	90	0	71.5
C12	32	90	100	48	0	67.5
C13	6	6	69	43	40	32.8
C14	11	85	0	69	0	33.0
C15	90	43	11	58	53	51.1
C16	100	64	95	79	47	76.9
C17	53	11	85	6	67	44.3
C18	69	79	6	22	13	37.9
C19	74	22	16	95	0	51.8
C20	0	100	53	16	87	51.1
C2V	77	70	58	100	100	80.8
C9V	37	100	100	100	0	67.3
C16V	100	77	100	95	56	85.6
C20V	0	100	64	14	100	55.4

rank-order of a company based on DEA and on KPIs average reflects the different focus of the two methods. DEA evaluates the DMUs allowing them to select their own weights, such that good performance in specific areas can be valued and differentiation strategies rewarded. KPIs average gives equal importance to all dimensions, rewarding balanced performance. Therefore, the two methods may be used as complementary tools in web benchmarking platforms. In our view, DEA enables obtaining rankings more aligned with company strategy, as the flexibility in the choice of weights enables reflecting

Table 5.11: DEA efficiency scores for the organizational perspective

DMU	Standard DEA Eff.	Restricted DEA Eff.	
	(*)	(*)	(**)
C9V	-	-	100%
C16V	-	-	100%
C2V	-	-	98.8%
C20V	-	-	98.3%
C16	100%	100%	95.3%
C9	100%	100%	94.5%
C20	100%	100%	93.8%
C2	100%	100%	88.7%
C12	100%	95.7%	87.3%
C3	100%	93.4%	86.8%
C11	100%	92.3%	82.0%
C7	100%	88.2%	77.4%
C15	94.2%	80.4%	76.0%
C8	100%	81.2%	74.5%
C17	100%	77.4%	71.1%
C18	96.8%	77.8%	71.0%
C14	89.0%	74.9%	70.5%
C19	98.5%	69.2%	64.8%
C6	76.9%	65.9%	61.1%
C4	81.8%	66.5%	60.6%
C13	75.3%	59.2%	54.8%
C10	64.9%	59.7%	54.8%
C1	44.7%	38.2%	35.4%
C5	64.4%	38.4%	35.1%

adequately decision maker preferences.

With the restricted composite indicator model, only four companies were considered efficient. These companies achieved the highest benchmark score in at least one KPI dimension, with the exception of C9 that achieved the second position in one KPI dimension. However, in terms of overall performance, C9 is in the first position of the KPIs average ranking. In particular, C16 was the best in Productivity and C2 was the best in both Accidents Frequency Rate and Sales Growth. It is interesting to note that C20 is the best in Profitability and also achieves a high level of Sales Growth (87%), but it is the worst in terms of Productivity. In these cases, where companies specialize on a particular set of dimensions outperforming the others, the DEA assessment has the ability to appreciate the differentiation strategy. This is an advantage of using DEA over other methods, since companies may legitimately wish to give more importance to certain aspects than to others. The Spearman rank-order correlation between restricted DEA results and KPIs average is equal to 0.882 (p-value equal to 0.000). It can be concluded that adding weight constraints to the DEA model increased the discrimination between efficient and inefficient companies and it also increased the correlation between DEA scores and KPIs average.

The P-DEA model was used to create virtual DMUs, C2V, C9V, C16V and C20V, based on the four efficient companies identified by the restricted DEA model (C2, C9, C16 and C20). Table 5.10 reports in the last lines the outputs of these virtual DMUs. Using the restricted composite indicator model with the virtual units, only two virtual companies were classified as efficient (C9V and C16V), as shown in Table 5.11. C9V is the best company in Profitability and Accidents Frequency Rate, C16V is the best in Productivity, and both achieved the highest score in Hanging Invoice. These two virtual companies define the new efficient frontier and should be regarded as best practice peers by all other companies. None of the observed companies was able to achieve a 100% efficiency score with this model, so this analysis is able to suggest directions for improvement even for the companies previously considered efficient (C2, C9, C16 and C20).

The DEA analysis also enables comparing the inefficient DMUs with their peers pointing directions for performance improvement. This managerial information is very important to drive company strategy, since the aspects that need more attention can be identified, and a detailed analysis of peers' practices may show how companies can improve their activity. This is illustrated using one of the worst performing companies C18, with an efficiency score equal to 71.0%. The values observed, targets and peers for C18 are shown in Table 5.12. The targets for each indicator correspond to the linear combination of the values observed in the peers, using the coefficients (λ_j) presented in Table 5.12. The peers for company C18 are C9V and C16V. For example, the productivity target for company C18 is 73, and is obtained as $(37 \times \lambda_{C9V} + 100 \times \lambda_{C16V})$.

Table 5.12: Targets and peers for company C18

	Company C18		**Peers Values** (%)	
Variables	Benchmark Values (%)	Targets (%)	C9V ($\lambda_{C9V} = 0.43$)	C16V ($\lambda_{C16V} = 0.57$)
PROD	69	73	37	100
PROF	79	87	100	77
HI	6	100	100	100
AFR	22	97	100	95
SG	13	32	0	56

From Table 5.12 it is possible to observe that C9V has better performance than C18 on Profitability, Hanging Invoice and Accidents Frequency Rate. Company C16V is better than C18 in all indicators with the exception of Profitability. Regarding the targets proposed for each indicator, it can be concluded that C18 has to increase its performance mainly in two dimensions: Hanging Invoice and Accidents Frequency Rate.

Note that the targets suggested are very demanding, as we imposed an improvement to all companies according to the standards defined by the practical frontier based on virtual DMUs. If companies which to follow a gradual path towards excellence, they can start by looking at the targets from the standard composite indicator model (5.1) without weight restrictions, and use the original DEA peers as examples to follow. In the case of company C18 the original peers were C9, C11 and C20. Figure 5.3 compares the KPIs values of C18 with the initial targets from the standard composite indicator model (5.1) and with the targets corresponding to the evaluation against the practical frontier.

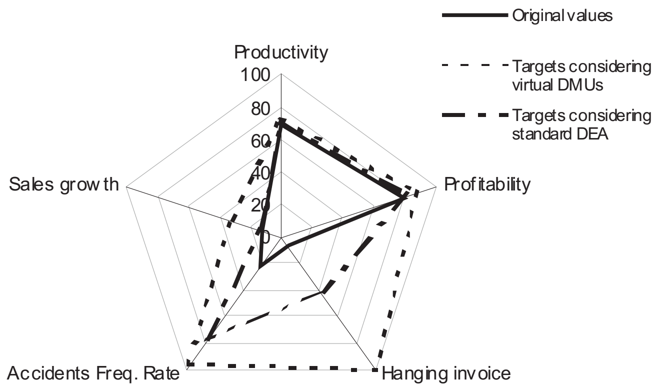

Figure 5.3: Targets for company C18 to follow a gradual path

5.4.2 Operations performance assessment

A similar analysis to the one reported in the previous section could be done for the operations model. The individual values associated to each operation can be seen in Table 5.13. Table 5.13 reports the benchmark data, the respective KPIs average and the efficiency scores. The symbol (∗) indicates an assessment without virtual operations, and the symbol (∗∗) an assessment including virtual operations. Table 5.14 reports the average DEA scores for the operations associated to each company.

Table 5.13: Operations benchmark data, respective KPIs average and efficiency scores

DMU	Operation #	CS-CC	CS-PA	CS-CW	CP	KPIs average	Standard DEA Eff. (*)	Restricted DEA Eff. (*)	Restricted DEA Eff. (**)
C2	O101	10	18	20	86	33.7	86.8	34.5	32.2
C4	O201	10	0	20	34	16.1	34.3	21.4	20.0
	O202	10	18	7	59	23.6	59.5	26.8	24.3
	O203	10	7	20	37	18.7	37.7	22.8	21.4
	O204	10	18	7	75	27.4	74.8	30.1	27.4
	O205	10	18	7	49	21.0	49.4	24.5	22.3
	O206	39	93	83	29	61.0	92.9	85.3	80.1
	O207	39	93	83	71	71.6	97.0	91.3	84.4
	O208	39	93	83	73	72.0	97.2	91.6	84.5
	O209	39	93	83	58	68.2	95.1	89.2	83.0
	O210	39	93	83	68	70.7	96.5	90.8	84.0
	O211	39	93	83	48	65.6	93.7	87.7	82.0
	O212	39	93	83	56	67.7	94.8	89.0	82.8
	O213	39	93	83	24	59.7	92.9	84.8	79.6
	O214	39	93	83	25	60.1	92.9	85.0	79.8
	O215	39	93	83	80	73.7	98.2	92.6	85.2
	O216	39	93	83	78	73.3	98.0	92.4	85.0
	O217	39	29	12	76	39.0	77.1	46.3	42.7
	O218	39	2	12	70	30.6	70.3	41.9	38.7
	O219	39	29	12	15	23.7	39.0	34.3	32.9
	O220	39	48	46	42	43.9	51.3	50.6	46.5
	O221	39	61	20	97	54.2	97.4	67.9	61.7
	O222	39	61	20	88	52.0	89.6	66.0	60.0
	O223	61	29	46	90	56.3	91.0	67.1	62.0
	O224	61	18	46	17	35.4	61.0	53.0	50.8
	O225	61	7	46	19	33.1	61.0	52.1	49.9
	O226	61	48	46	9	40.9	61.0	55.3	52.9
	O227	61	7	7	95	42.5	96.1	61.7	57.0
	O228	86	29	83	98	74.1	100.0	89.5	82.7
	O229	86	61	83	85	78.7	97.2	90.1	83.3
	O230	86	36	83	14	54.7	86.4	77.0	73.7
	O231	86	29	98	0	53.3	98.3	84.5	80.3
	O232	86	48	83	53	67.6	86.4	82.4	78.9
	O233	86	36	83	93	74.6	98.4	89.2	82.4
	O234	86	93	46	41	66.5	93.0	87.7	82.3
	O235	86	93	46	3	57.1	92.9	83.7	78.6
	O236	86	36	83	20	56.4	86.4	77.7	74.4
	O237	100	100	100	39	84.8	100.0	100.0	93.9
	O237V	100	100	100	51	87.8	-	-	95.1
	O238	100	93	98	31	80.4	100.0	96.3	92.2
	O239	100	61	83	12	63.9	100.0	89.4	85.6
	O240	100	93	98	63	88.5	100.0	99.6	95.4
	O241	100	61	83	2	61.4	100.0	88.3	84.6
	O242	100	29	98	5	58.0	100.0	86.9	83.2
C7	O301	61	18	46	46	42.6	62.8	56.5	53.7
	O302	61	48	46	81	59.1	85.7	67.4	62.3
	O303	61	100	46	44	62.7	100.0	90.6	85.1
C13	O401	61	0	46	10	39.0	61.0	50.5	48.3
	O402	86	0	98	54	79.6	98.9	87.2	82.9
	O403	86	0	98	64	83.0	99.3	88.3	83.9
C14	O501	86	61	83	61	72.8	87.7	84.9	81.0
C15	O601	0	100	0	100	50.0	100.0	88.1	80.0
	O602	61	100	46	22	57.2	100.0	88.3	82.9
C17	O701	86	48	83	32	62.5	86.4	80.3	76.8
C18	O801	61	61	46	7	43.6	61.0	57.5	54.0
	O802	61	48	46	27	45.5	61.0	57.3	54.8
	O803	61	36	46	36	44.5	61.0	56.8	54.4
	O804	86	93	98	83	90.2	100.0	100.0	95.0
	O804V	100	100	100	100	100.0	-	-	100.0
	O805	86	48	83	92	77.3	99.0	90.2	83.3
	O806	100	93	98	66	89.3	100.0	100.0	95.7
	O806V	100	100	100	86	96.5	-	-	98.6
	O807	100	93	98	51	85.5	100.0	98.4	94.2

Table 5.14: DEA average efficiency scores for the operations perspective

DMU	No. operations	Standard DEA average eff.	Restricted DEA average eff.	
		(∗)	(∗)	(∗∗)
C15	2	100%	88.20%	81.40%
C14	1	87.70%	84.90%	81.00%
C2	1	86.80%	34.50%	32.20%
C17	1	86.40%	80.30%	76.80%
C13	3	86.40%	75.30%	71.70%
C4	42 +1 virtual	84.30%	72.00%	68.10%
C18	7 + 2 virtual	83.10%	80.00%	81.10%
C7	3	82.80%	71.50%	67.00%

With the standard composite indicator model, twelve operations achieved 100% efficiency, but only C15 achieved 100% efficiency for all its operations. Including weight restrictions, only three of the twelve operations previously classified as efficient obtained a 100% efficiency score. Two of the efficient operations are associated with C18 and one is from C4. In this case, none of the companies had all its operations classified as efficient.

Similarly to the organizational model, adding weight constraints increases the discrimination power of the analysis. In addition, the Spearman rank-order correlation between DEA efficiency scores and KPIs average increases from 0.786 (for the standard composite indicator model) to 0.934 (for the restricted composite indicator model).

The three efficient operations were used to specify the virtual DMUs used in the P-DEA model. Including the virtual operations in the weight restricted DEA assessment only one virtual operation remained efficient and all the other operations decreased their scores, as expected. This virtual DMU represents the best practice standard.

5.4.3 Confrontation of the two perspectives

This section compares the results of the two performance assessment perspectives based on the weight restricted DEA model including both the original and the virtual DMUs. This allows obtaining a performance overview of the companies, including financial and operations aspects. Figure 5.4 shows the position of each company in the two perspectives: organizational efficiency (vertical axis) and operations efficiency (horizontal axis). Since the operations model includes only eight companies, this comparison is limited to this subset of the original sample.

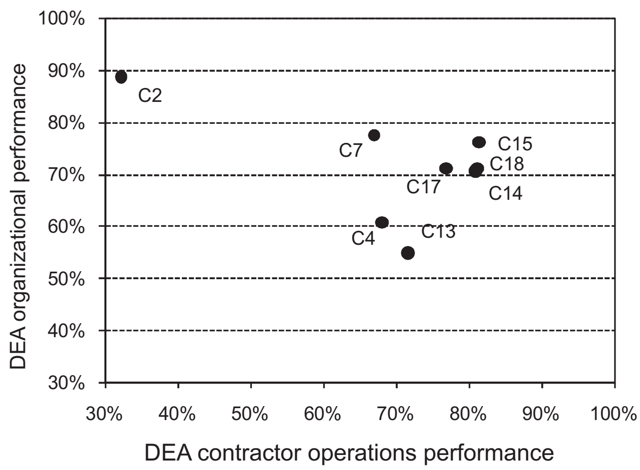

Figure 5.4: Companies efficiency regarding the two perspectives

From Figure 5.4, it is possible to see that most companies achieved similar performance levels in both dimensions. However, it is interesting to observe that company C2 is the best performer on organizational efficiency but it shows weaknesses in terms of operational efficiency, as it is the worst company in this dimension. Although these results cannot be extrapolated to Portuguese contractors in general, this can be an indication that some companies focus their efforts in certain aspects like, for instance, to achieve better financial results, possibility neglecting other aspects considered less relevant to pursuit the company strategy.

Note that the efficiency values produced by DEA are relative, so they are valid for a particular sample. A contractor that is efficient compared with a given sample may be quite inefficient when compared with another sample. In other words, if a group of very poor contractors were evaluated using DEA, efficient contractors would still be identified. This implies that even for a company with a high efficiency score there might be potential for improvement. This can be identified with the analysis of individual KPIs, where low values suggest there is scope for improvement. For example, company C2 achieved a high efficiency score in the organizational performance perspective. However, the analysis of KPIs suggests that it may be possible to improve in Productivity (PROD=64%), Profitability (PROF=58%) and Hanging Invoice (HI=48%), although there is no evidence in the sample that it can improve all dimensions simultaneously.

5.5 Summary and conclusions

The analysis reported in this study primarily intended to develop a methodology to complement the information provided by a set of KPIs available in web benchmarking platforms. A DEA analysis is a useful method to obtain a single overall performance measure for each company, and to provide important managerial insights concerning company rankings and targets for performance improvement.

The fact that the majority of CI companies are small or even very small organizations gives an extra importance to the development of easy to use tools for management decision and strategy, since most of them have their senior staff concentrated in practical action, with little time to dedicate to a deeper analysis of performance assessment results and of what lies underneath. Therefore, this study is a contribution to assist companies involved in benchmarking experiences aiming to improve their effectiveness.

This chapter assessed the efficiency of a sample of Portuguese contractors representing some of the leading companies operating in the sector in terms of value of sales and number of employees. The contractors were assessed from two different perspectives: organizational performance and operations' performance. Their combination provides an overview of company performance. To obtain a more realistic company assessment, two types of DEA models were used in each perspective. The standard DEA model allows the weights assigned to each of the indicators to vary freely, such that all companies appear at their best. The weight restricted DEA model was used to incorporate experts' opinion concerning the relative importance of the indicators considered. To enable defining targets for all companies, including those originally considered efficient, the restricted DEA model was run with a sample including virtual organizations to explore the advantages of assessing performance against a practical frontier whose standards represent an improvement in relation to the productivity levels currently observed.

Future research will attempt to increase the size of the sample analyzed in order to accomplish a better representation of contractors. In addition, future analysis should attempt to include in the models data related to deviations from the project due date and indicators of client satisfaction, as these aspects are critical in this sector. This requires the implementation of procedures that enable consistent data collection for a significant number of companies. Finally, this quantitative analysis of contractors should be followed by a qualitative research consisting on visits to benchmark companies in order to identify best practices and spread them in the sector.

Chapter 6

Performance and innovation in the Portuguese construction industry: a study of factors promoting financial soundness and innovation in the industry

6.1 Introduction

The Construction Industry (CI) is one of the largest sectors in Portugal. It represents 10.7% of total employment and 6.4% of Gross Domestic Product (GDP). The majority of companies in the Portuguese CI are micro or small companies. In particular, in 2011 the number of general contractors operating in the CI is 23585, with the large companies (more than 100 employees) representing less than 1.2% of general contractors (www.inci.pt).

The Portuguese construction industry has been witnessing substantial changes over the years. After joining the European Union in 1986, Portugal benefited from structural funds that promoted the development of infrastructures, such as the major national roads network, bridges and social facilities. The fall of the nominal interest rates in 1995 contributed to the expansion of the residential housing segment. Due to the joint effect of these factors, the Portuguese construction industry experienced a remarkable development during the 1990's. In particular, new Portuguese construction companies were created, and a few international companies, mainly from Spain, expanded their business to Portugal. In order to cope with the level of demand, Portuguese companies adopted strategies of subcontracting specific stages of the productive process. Subcontracting became a common practice in the sector, which caused the downsizing of some companies that specialized in particular skills.

The growing trend in the construction activity started to invert in 2001. The Portuguese economy was hit by a financial crisis, and the public investment was reduced. As a consequence, the Portuguese construction activity slowed down. After 2007 the contextual setting got worse due to the slump in the global economy. Since then, the construction industry emerged in a period of downturn. Portuguese companies had to redirect their strategies to ensure their viability in the market. Some companies expanded their operations to external markets, namely Latin America, Eastern Europe and Africa, and others adopted strategies of business diversification into activities related to construction, e.g. rehabilitation or renewable energies. Nowadays, the industry is regarded as being oversized for the current and prospective needs of the country. With such an adverse environment, it becomes vital to examine the performance of the Portuguese construction companies in order to promote sustainable efficiency improvements and to encourage excellence in the sector.

The purpose of this chapter is to evaluate the financial performance in the Portuguese CI in the recent past, between 1996 and 2009, identifying the factors that promote excellence and innovation in the sector. In addition, this chapter proposes an enhanced methodology to assess innovation within an industry, identifying the innovative companies and quantifying the degree of innovation.

The performance of the construction companies was evaluated with a composite indicator estimated using DEA. The individual indicators underlying the construction of the composite indicator characterize company financial soundness, and are related to profitability, financial autonomy, liquidity and value added. DEA is a deterministic approach, such that the evaluation of performance assumes that there are no random factors affecting the construction of the frontier. This is a shortcoming of DEA, as there may be random noise or measurement errors in the data. To overcome this limitation, the DEA evaluation was complemented with bootstrapping to ensure a robust evaluation of performance, and enable statistical inference.

In a second stage of the analysis reported in this study, the determinants of good performance and innovation were explored using regression techniques. In particular, we explored the relationship between performance and company size, headquarters location, engagement on Research and Development (R&D), and the national economic environment.

The remainder of this chapter is organized as follows. Section 6.2 describes the methodology used in this chapter in greater detail. Section 6.3 includes the data used in the study, and Section 6.4 discusses the results obtained. The last section concludes and suggests recommendations for future research.

6.2 Methodology

6.2.1 Performance assessment

The methodology adopted in this chapter involves the estimation of Portuguese construction companies financial performance with a composite indicator estimated using DEA. We run model (5.1) including a pooled sample of the Portuguese CI companies, comparing the financial indicators of each company in each year with a frontier representing the best practices of all years analyzed. From the results of model (5.1), the average value of the composite indicator for the companies analyzed in a given year provides an idea of the average industry performance of that year in relation to the pooled frontier. For example, a low average value can be interpreted as an indication that the industry had poor financial performance during that year, whereas a high average value suggests that the companies performance was close to the best-practices observed in the period studied. Furthermore, a large dispersion of performance scores in a given year suggests that good practices co-exist with poor financial performance, whereas a small range of scores indicates that companies performance is homogeneous.

After computing the composite indicator, we used bootstrapping to obtain unbiased performance estimates. The bootstrapping analysis was implemented using the statistical package R including the FEAR library, developed by Wilson (2008).

6.2.2 Explaining performance

The impact of management practices and exogenous factors on performance only recently started to be explored in the CI literature. The studies developed to date focused on the impact of best practices on company or project performance. For instance, El-Mashaleh et al. (2006) analyzed the impact of using information technologies on company performance. For a literature review on the factors that may impact project performance see Korde (2005).

Until recently the common approach to regress DEA efficiency scores against explanatory variables was to employ a tobit regression. A number of authors argued that tobit regression may be appropriate since the efficiency values are bunched at one. However, the tobit regression is appropriate for censored data, which is not the case of DEA efficiency scores, as there are no estimates greater than one that need to be censored at one to undertake a regression. Simar and Wilson (2007) argued that the regression model that should be estimated in a second stage assessment should include a truncated term rather than a censored term. As the DEA efficiency estimates can be biased, and serially correlated in a complicated unknown way, Simar and Wilson (2007) demonstrate that a second stage procedure requires correcting for bias the efficiency scores and correcting for correlation. In particular, the authors suggest two algorithms for second stage analysis. As presented in Simar and Wilson (2007), algorithm #1 proposes to use the original efficiency (not corrected by bias) in the truncated regression, and bootstrap truncated regression (single bootstrap). Algorithm #2 proposes to use the bias-corrected efficiency in the truncated regression and bootstrap truncated regression (double bootstrap). The authors demonstrate that the double bootstrap is the most suitable approach.

To explore the factors that may be associated with good performance levels of the construction companies, we used a truncated regression with bootstrapping, specified according to the algorithm #2 proposed by Simar and Wilson (2007). The model specified for our analysis was formulated using company size, headquarter location, engagement on R&D projects, and national GDP per capita (as a proxy of the national economic context of each year) as regressors.

Company size was included in the model to explore the existence of economies of scale in the CI. The issue of whether larger companies have superior performance compared to smaller companies, or vice-versa, generated a large amount of theoretical and empirical research in many activity sectors, such as in manufacturing (Soderbom, 2004; Halkos and Tzeremes, 2007), transportation (Odeck, 2008), or banking (Ray, 2007). As explained by Fiegenbaum and Karnani (1991) larger companies may have the advantage of strategic diversification, negotiating power, and means to face competition by keeping prices below the competitive level. According to Strigler (1939), small companies are able to compete successfully with large, more static-efficient producers, by using more flexible production technologies. Smaller companies have the advantage of adjusting key competencies to exogenous shocks at relatively low costs. In the CI there are just a few studies that analyze the relationship between company size and performance. For instance, Kale and Arditi (1998) concluded that small company size was one of the major factors to explain business failure. The study developed by Maes et al. (2005) revealed that the company size has no significant direct impact on financial performance. This topic is still controversial in the CI, but essential to redirect strategies to improve performance.

Geographic location was included in the model to test whether construction companies located in Lisbon and its vicinity are more efficient that the others. According to a report produced by AEP (2007), the Portuguese construction companies are mainly located in the North (27% of the companies), Center (31% of the companies), and Lisbon area (25% of the companies). Although the degree of regional concentration is relatively small in Portugal, it would be interesting to test if companies with headquarters located in the capital area may benefit from being close to the dynamic activity that typically characterizes capital cities.

The model included R&D engagement to test its relationship with the performance of construction companies. The studies by Arditi and Gutierrez (1991) and Someya (1992) analyzed this issue in the context of United States construction companies. Both studies concluded that the R&D activity tends to affect United States contractors' performance. It is generally believed that CI low productivity may be due to the lack of commitment to R&D. The study described in this chapter provides further empirical evidence on the relationship between R&D and company performance. The engagement in R&D is measured as a dummy variable indicating if a company undertook, or not, R&D projects internally or in cooperation with academic institutions. As the data available was scarce, it did not allow a more detailed quantification of companies' R&D engagement, so the results obtained from this analysis should be interpreted with caution.

The economic context was included in the model to check its influence on the performance of the CI. The CI is a cyclical sensitive industry (Moscarini and Postel-Vinay, 2009), whose dynamics is often considered one of the main indicators of the economy or one of its barometers. In particular, as mentioned by Ngowi et al. (2005), the CI demand can be determined by investment or the stage of economic development and growth. Although it is broadly assumed that the economic context has an impact on the performance of construction companies, only a few studies (e.g. Akinsola et al., 1997) tested the significance of the relationship between the national economic context and company performance.

Concerning the specification of the truncated regression model, due to the availability of data from different time periods, we used a panel data truncated model controlling both for company and time effects. The model is specified as follows:

$$C_{jt} = \alpha_o + \delta_j + \eta_t + z_{jt}\beta + \varepsilon_{jt} \tag{6.1}$$

Subscript j represents the j^{th} company ($j = 1, \ldots, n$), subscript t represents the time period ($t = 1, \ldots, l$), α_o is an intercept, δ_j is a vector of dummy variables for each company, η_t is a vector of dummy variables for each year, z_{jt} represents the set of regressors previously identified, β denotes the regression coefficients and ε_{jt} is the error term with a $N(0, \sigma_\varepsilon^2)$ distribution with a truncation at $(1 - \alpha_o - \delta_j - \eta_t - z_{jt}\beta)$. Note that C_{jt} corresponds to the performance level of company j in year t, estimated using model (5.1) and corrected by bootstrapping.

Concerning the bootstrapping on truncated regression, the algorithm involves the following steps:

1. Compute the performance estimate $\widehat{E_j}$ for each DMU $j = 1, \ldots, n$.
2. Estimate using maximum likelihood the truncated regression of $\widehat{E_j}$ on z_j, yielding $(\widehat{\beta}, \widehat{\sigma}_\varepsilon)$.
3. Compute B ($B = 2000$) bootstrap estimates for β and σ_ε in the following way:
   - for $j = 1, \ldots, n$ draw ε_j from $N(0, \widehat{\sigma}_\varepsilon^2)$ with a truncation at $(1 - z_j\widehat{\beta})$;
   - for $j = 1, \ldots, n$ compute $E_j^* = z_j\widehat{\beta} + \varepsilon_j$;
   - estimate using maximum likelihood the truncated regression of E_j^* on z_j, yielding estimates $(\widehat{\beta}^*, \widehat{\sigma}_\varepsilon^*)$.

6.2.3 Assessment of innovation

The final stage of the performance assessment is related to the evaluation of innovation. In this study, a company is considered an innovator if it shifts the best-practice frontier of a given year to better levels than those observed in previous time periods. This means that an innovator company is able to introduce better practices in the industry. The other companies (i.e. the followers, hereafter) can improve their performance by copying the practices observed in the innovators.

The assessment of innovation using DEA models was originally proposed by Fare et al. (1994a), and served as the basis for the analysis described here. In this study, the concept of innovation was adapted to an evaluation of performance based on composite indicators. Our approach concerns the comparison of the achievements of a company in year t with the practices observed in all previous years, up to year t, whereas the concept of innovation proposed by Fare et al. (1994a) only involved a comparison of practices between two consecutive years.

To be able to identify the innovators, we run a linear programming model comparing the performance of each company, in a given year t, with the performance of all other companies, including observations from the same year and from previous years. As proposed by Tulkens and Vanden Eeckaut (1995), a frontier that results from an assessment whose production possibility set is defined including observations from a given year as well as from previous periods is called a sequential frontier. As presented in Thiry and Tulkens (1992), this approach is based on the idea that what was feasible in the past remains achievable in any later period. We consider that a necessary condition to be innovator in year t is to be located in the best practice sequential frontier, which implies having a composite indicator equal to one in a sequential frontier assessment.

To quantify the magnitude of the performance enhancement of innovative companies in relation to previous years, we run a second linear programming model comparing the performance of the innovators in year t, with all other companies in previous years, up to $t-1$. The estimate obtained using this procedure is called innovation score hereafter. Note that this second linear programming model differs from the previous assessment due to the exclusion of observations from year t. The innovation score can either be equal to one, meaning that the achievement of company j in year t is identical to what was observed in previous years, or greater than one, meaning that the company has actually moved the frontier to more productive levels than those previously observed.

The final stage of the analysis of innovation consisted on the study of the factors that promote innovation in the CI. For this purpose, we constructed a logistic regression model, where the dependent variable characterizes the companies as innovators or followers. The logistic model is used because the dependent variable is binary. The factors used to predict innovation were the average company size, geographic location of the headquarter, and engagement on R&D projects. We also include the average value of profitability and value added for each company to verify if the financial status of the company and labor productivity have an impact on innovation. In the model proposed, a value of one for the dependent variable represents the companies that innovated in at least one of the years analyzed, and zero represents the remaining companies - the followers. The logistic model predicts the odds of being innovator, given known values of the independent variables. The odds is defined as the ratio of the probability of being innovator to the probability of being follower. The logistic model can be specified as follows:

$$ln\left(\frac{P_j}{1-P_j}\right) = \alpha_o + z_j\beta, \tag{6.2}$$

or in terms of the predicted probability of being innovator as:

$$P_j = \frac{1}{1+e^{-(\alpha_o+z_j\beta)}} \tag{6.3}$$

In the expressions above, P_j is the predicted probability of being innovator for the j^{th} company, and z_j represents the set of regressors: company size, geographic location, R&D engagement, value added and profitability. Companies with P_j values above 0.5 were classified in the innovators group, and companies with P_j values below 0.5 are classified in the followers group. The classification accuracy of the model is obtained by comparing the actual with the predicted status.

6.3 The data

This chapter uses a sample of 110 companies corresponding to major Portuguese contractors laboring on public works. The longitudinal assessment reported covers the time period between 1996 and 2009. The data used came from the database of the 500 larger Portuguese companies, in terms of value of sales, published every year in the *Exame* magazine (www.exame.pt). The construction companies sampled vary from one year to the other, which is explained by the volatility of the market. In particular, new companies enter the market, and others cease activity, merge or do not provide data in some years. This implies that the assessment explores an unbalanced panel data sample.

In this study we used a performance assessment model that characterizes contractors from a financial perspective. The indicators selected for a financial benchmarking model should capture the multidimensional nature of the CI activity, and the overall company performance. Based on the review of CI literature (Pilateris and McCabe, 2003; Severson et al., 1994; Kangari et al., 1992; Curtin, 1993; Altman, 1968), it can be concluded that the most critical financial ratios to measure contractor performance include: profitability, leverage and liquidity. Therefore, the performance assessment model specified includes indicators representing the three critical financial ratios: Profitability (PROF), Financial Autonomy (FA), and Liquidity (L) as well as another indicator related to the value added of the company's employees (VA). This variable is particulary important in the CI because it highly depends on the labor force to be profitable. The indicators used in the model are described in Table 6.1. In terms of indicators' definition, Profitability measures the profit of the company before tax and interest. Value Added measures the contribution to Gross Value Added made by an individual employee. The VA indicator was deflated using the GDP deflator, available in the World Bank's World Development Indicators database (http://data.worldbank.org/), considering 1996 as the base year. Financial Autonomy measures the contribution of equity on company funding to provide an indication of the long-term solvency. Liquidity measures the company ability to meet short-term financial obligations. These last two indicators have minimum obligatory requirements to allow contractor activity in Portugal (at least 100% on Liquidity and 5% on Financial Autonomy, since the year of 2010).

The DMUs assessed correspond to construction companies, characterized by four KPIs, all of which are intended to be maximized. Note that these KPIs correspond to the output indicators used in model (5.1).

Table 6.1: Indicators and respective formula

Indicators	Formulas	
Profitability	$PROF = \frac{P}{V} \times 100$	P-Profit before taxes (€) V-Value of sales(€)
Value Added	$VA = \frac{GVA}{N}$	GVA-Gross value added (10^3€) N-No. of employees
F. Autonomy	$FA = \frac{E}{A} \times 100$	E-Equity (€) A-Value of assets (€)
Liquidity	$L = \frac{FA}{FL} \times 100$	FL-Floating liabilities (€) FA-Floating assets (€)

Before proceeding to frontier estimation, we explored the existence of outliers in the sample. It is known that in the presence of outliers, the location of the DEA frontier may be severely affected due to its sensibility to extreme observations. The order-m method suggested by Cazals et al. (2002) was used to mitigate the impact of potential outlier behavior. The order-m approach is based on the concept of expected maximum output function (or minimum input function) yielding frontiers of varying degrees of robustness. In particular, the semi-automatic procedure proposed by Simar (2003) for outlier detection was applied. Table 6.2 reports the number of observations included in the analysis and the descriptive statistics of the variables used in the DEA model. Table 6.2 also reports the number of observations that were considered outliers in each year, and consequently excluded from the assessment. From the original 110 companies, 13 were considered outliers in all years with data available. Therefore, the final sample comprised 97 companies. Most of these companies did not provide data for some of the 14 years considered (from 1996 to 2009), so the total number of observations considered in the performance assessment was 567.

From Table 6.2, it is possible to observe that the companies analyzed are relatively homogeneous, given the small values of standard deviation. Profitability is the indicator that reveals the greatest variation across years.

Table 6.2: Descriptive statistics of data

Year	No. obs.	No. outliers	Liquidity	F. Aut.	V. Added	Prof.
1996	29	7	139.8	21.3	26.0	1.6
1997	35	9	133.5	22.2	25.5	2.0
1998	35	6	147.3	21.9	24.4	1.5
1999	35	6	146.0	25.9	25.5	2.0
2000	38	8	151.7	25.2	28.8	2.6
2001	41	8	152.9	24.5	29.7	2.4
2002	48	5	161.2	26.0	29.3	2.6
2003	56	5	165.6	27.8	30.2	3.1
2004	46	8	157.3	26.3	33.0	2.9
2005	47	9	157.5	26.0	36.9	3.2
2006	52	7	159.0	26.7	37.2	3.2
2007	32	4	170.3	25.3	32.1	2.9
2008	36	4	147.5	21.1	31.6	2.6
2009	37	6	151.9	22.1	33.3	3.5
		Mean	154.0	24.7	30.7	2.7
		St. dev.	46.1	8.0	12.0	2.4

6.4 Results and discussion

6.4.1 Performance assessment

The first stage of the assessment was intended to explore whether the performance level of Portuguese construction companies improved over time. This required the estimation of a composite performance indicator for each company in each year, which was based on a comparison with a pooled frontier representing the best practices observed in the 14 years analyzed.

Table 6.3 reports the summary results for the composite indicator obtained using model (5.1). The bootstrap results are also shown in Table 6.3, including the average of the bias-corrected estimates, bias, standard deviation, and 95% confidence intervals for the bias-corrected composite indicator in the corresponding year.

Table 6.3: Results of original and bootstrapped performance estimates

Year	Orig. Eff	Bias-corr Eff	Bias	St. dev	Lower bound	Upper bound
1996	0.555	0.530	0.025	0.011	0.514	0.549
1997	0.563	0.533	0.030	0.013	0.516	0.555
1998	0.591	0.560	0.031	0.014	0.543	0.584
1999	0.618	0.588	0.030	0.013	0.571	0.611
2000	0.636	0.605	0.031	0.014	0.587	0.628
2001	0.617	0.588	0.029	0.013	0.570	0.610
2002	0.650	0.617	0.033	0.014	0.598	0.641
2003	0.694	0.653	0.041	0.018	0.633	0.684
2004	0.682	0.642	0.039	0.017	0.623	0.673
2005	0.697	0.648	0.049	0.021	0.629	0.686
2006	0.705	0.650	0.055	0.022	0.633	0.694
2007	0.678	0.639	0.039	0.016	0.619	0.668
2008	0.606	0.569	0.037	0.016	0.551	0.598
2009	0.635	0.587	0.048	0.021	0.568	0.624

From Table 6.3 we can conclude that from 1996 to 2000 the bias-corrected financial performance of the Portuguese construction companies increased approximately 14%, which is a remarkable improvement. It is interesting to observe that this period coincides with the booming of the Portuguese CI, when some of the major Portuguese public works were constructed, such as Expo'98 and Vasco da Gama bridge. From 2001 to 2009 the performance improved approximately 3%, which corresponds to a slight improvement. This result is as expected since after 2001 the Portuguese construction activity slowed down. The slightly improvement in performance during 2002 and 2003 may be explained due to the hosting of the EURO 2004 football championship that required the construction of several infrastructures and stadiums. In 2007 and mainly in 2008, it was observed an abrupt performance decline due to the severe global financial crisis that began in North America and immediately spread into Europe.

From the bootstrapped results, we can observe that the bias-corrected estimate is within relatively narrow confidence intervals, i.e. the lower and upper bounds of the intervals are relatively close. This provides statistical confidence for the bias-corrected estimate. Moreover, as the estimated bias is much larger than the standard deviation for all the companies, we conclude that the bias-corrected estimates should be preferred to the original estimates, as explained in Fried et al. (2008, pp.467). Thus, the bias-corrected estimates were used in the analysis reported hereafter. Analyzing the bootstrap confidence intervals, we can confirm the previous conclusions concerning performance changes over time. Although the confidence intervals from consecutive years overlap, it is possible to confirm the significant performance improvement from 1996 to 2000, and that the performance did not change significantly from 2001 to 2009.

To complement the performance assessment of Portuguese construction companies, we investigated whether good and bad performers co-existed in each of the years. To explore this issue, we plotted the density of bias-corrected performance estimates for each year. We used the univariate kernel smoothing (Wand and Jones, 1995), and the reflection method to determine densities for the performance estimates. The criterion for bandwidth selection followed the plug-in method proposed by Sheater and Jones (1991). Figure 6.1 plots the estimated densities of the bias-corrected performance distribution for each year.

From Figure 6.1 we can observe that, overall, the shape of the distribution did not change significantly as the distributions of the bias-corrected efficiency scores are always unimodal. The years which have a higher probability density on higher levels of performance correspond to 1997, 1998, 2003, 2006, and 2007, indicating that the number of good performers observed in these years is higher than in the remaining years. It is interesting to note that most of these years (1997, 1998, 2006, 2007) coincide with the years

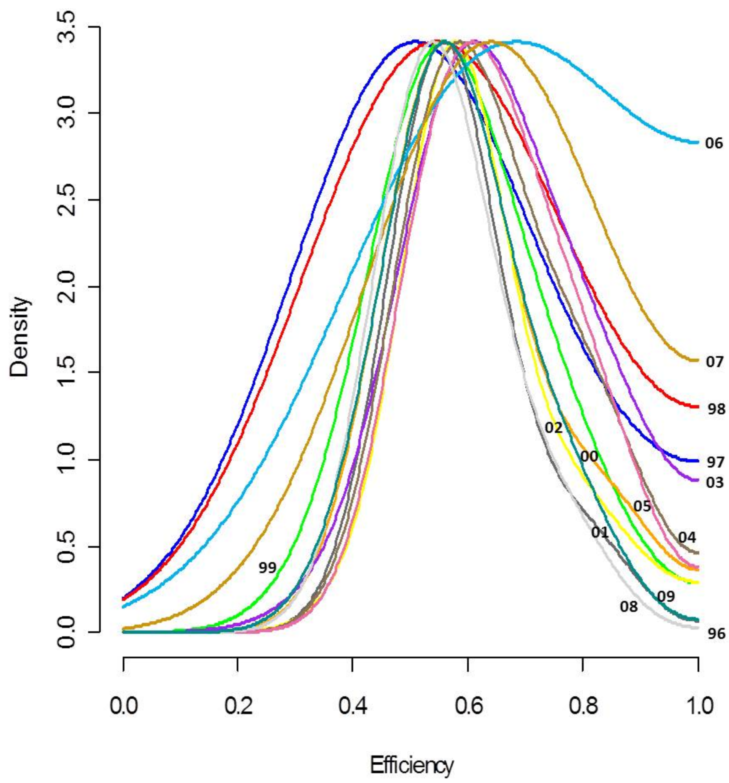

Figure 6.1: Non-parametric kernel estimates of bias-corrected scores

with a high number of bad performers as well. This means that in these years the gap between good and bad performers was higher, which indicate that companies were more heterogenous in terms of financial performance.

6.4.2 Drivers of good performance

The purpose of this section is to explore the determinants of good performance. In particular, we explored the impact of the economic context, headquarter geographic location, company size and engagement on R&D projects on company financial performance. The regression model specified is a panel data truncated model, with the bias corrected composite indicator used as dependent variable, and the four factors mentioned above as independent variables.

The national economic context was characterized by GDP per capita, measured in US dollars. The data was taken from the World Bank's World Development Indicators database. This variable was deflated using the GDP deflator, considering 1996 as the base year. Table 6.4 reports the Portuguese GDP per capita, in real values.

Table 6.4: Portuguese GDP per capita in real values

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
GDP	12032	11016	11219	11126	9953	9791	10298	12152	13494	13558	13825	15405	16461	15215

The company size variable was measured by the value of sales, in real terms. The square of the value of sales was also included in the model to allow for the existence of variable returns to scale. The geographic location of companies corresponds to the location of headquarters. We considered a dummy variable distinguishing companies with headquarters located in the capital or its vicinity from the other companies. The engagement on R&D projects corresponds to a dummy variable distinguishing between "engaged" or "not engaged" on R&D projects internally or in cooperation with academic institutions. This qualitative information was provided by CI academic experts routinely involved in conducting R&D projects in many CI companies, with a deep knowledge of the Portuguese CI and its players. Table 6.5 details the geographic location and the number of companies engaged on R&D projects included in the sample.

Table 6.5: Number of companies by location and engagement on R&D

Company type	Lisbon area	Other locations
Engaged in R&D	17	20
Not Engaged in R&D	28	32

Table 6.6 reports the estimates from the panel data truncated model, the coefficients, standard errors, and p-values. The total number of observations

included in the model was 567, corresponding to all observations analyzed in all years. Note that we also used time and company dummies to control for time and company effects. The overall regression model was found to be statistically significant (χ^2 test with p-value of 0.000), with a pseudo-R^2 equal to 0.60.

Table 6.6: Regression analysis results

Variable	Coef.	Std. error	p-value
GDP	0.029	0.007	0.000
Size	-0.103	0.019	0.000
Size squared	0.009	0.003	0.001
Lisbon area	-0.332	0.099	0.510
R&D engagement	-0.033	0.088	0.711
Constant	0.411	0.135	0.002

From the results of Table 6.6, we can conclude that the national economic context has an impact on the performance of CI companies. This impact is positive, meaning that a higher GDP promotes better financial performance of the CI companies. This result confirms the importance of the national economic context to explain the performance level of the construction companies.

Company size is related to performance nonlinearly, as both the first and the second order coefficients are statistically significant. It is found a U-shaped relationship between performance and size. In particular, performance first decreases as company size increases, and for a value of sales greater than 585 million euros, performance starts to increase as size further increases. This means that the best performance levels tend to be achieved by small specialized companies, as well as by large companies. This result could be expected due to the highly fragmented nature of the CI in Portugal.

The hypothesis that the engagement on R&D projects has a direct impact on financial performance was rejected. This finding is surprising but may be explained due to the limitations associated with the measurement of the R&D variable. The long-term and the quantification of the impact of R&D engagement on company performance remains an open question for further research.

The regression results show that the headquarter location is not influential, meaning that companies with their headquarters located in the capital or its surrounding area do not outperform the other companies. This could be explained due to the small dimension of Portugal, enabling a nationwide company activity.

6.4.3 Companies that innovated

This section explores innovation among CI companies in the period 1996 to 2009, following the procedure described in section 6.2.3. Innovation in the context of this study reflects the ability of the CI companies to shift the frontier of financial achievements to better levels than those observed in previous years.

Table 6.7 shows the companies that were considered innovators in at least one year. The years that each company achieved the status of innovator are signalled with the letter I. For the other years, the symbol − indicates that the company did not provide data, and H (or L) indicates that the performance score of the company was higher (or lower) than the industry average for that year. Table 6.7 also includes the number of companies considered innovative, and the average innovation score for each year analyzed.

This analysis revealed that 22 companies from the 97 included in the sample were considered innovators in at least one year. Several companies (9) were considered innovative just in one year. Only 5 companies were considered innovators in more than half of the years for which data was available for their evaluation. The number of companies able to bring innovation to the

Table 6.7: Innovative companies characterization

Company	97	98	99	00	01	02	03	04	05	06	07	08	09	Total
Edificadora Luz Alves	I	H	I	I	I	H	I	-	-	-	I	-	-	6
Zagope	H	I	-	I	I	L	H	L	I	L	L	L	H	4
Lena Construções	H	I	I	I	I	-	H	H	H	-	H	H	L	4
H.C.I. Construções	-	-	-	I	-	-	I	H	I	H	-	-	H	3
Seth	-	-	-	-	-	I	I	-	-	I	-	-	-	3
Somague (Madeira)	-	-	-	-	-	-	I	H	I	I	-	-	-	3
Teixeira Duarte	I	-	-	-	-	H	I	H	-	H	H	L	L	2
Somafel	L	H	H	H	I	I	H	-	-	-	-	-	-	2
Epul	-	I	-	-	-	H	-	H	H	I	-	-	-	2
Casais	I	I	H	H	H	H	H	H	L	L	L	L	L	2
M.R.G.	-	-	-	-	-	-	I	I	H	H	H	-	-	2
Montiterras	-	-	-	-	-	-	-	I	-	I	-	-	-	2
Avelino F. Agrela	-	-	-	-	-	H	I	H	-	-	-	I	-	2
Bento Pedroso	H	H	H	-	-	I	-	H	H	-	-	-	-	1
Conduril	I	L	L	L	L	L	L	H	H	H	H	-	-	1
Termague	L	L	I	-	-	-	-	-	-	-	-	-	-	1
Tecnovia	L	H	-	I	-	L	L	L	H	L	L	L	-	1
Sopol	H	H	L	L	L	I	H	H	H	L	-	-	-	1
Pavia	-	H	H	I	H	H	L	L	-	-	-	-	-	1
Rosas Construções	L	-	-	-	L	H	L	L	H	I	-	-	H	1
Tecnovia (Madeira)	-	-	-	-	-	-	I	-	-	-	-	-	-	1
Tecnovia (Açores)	-	-	-	-	-	-	H	-	-	I	-	-	-	1
No. of innovators	4	4	3	6	4	4	8	2	3	6	1	1	0	
Mean innovation score	1.44	1.22	1.12	1.17	1.09	1.05	1.19	1.44	1.17	1.16	1.05	1.01	-	

industry in a given year varies between 0 and 8. This suggests that keeping in the cutting edge of innovation for long periods is a difficult task.

The years with the highest number of innovators are 2000, 2003, and 2006, which coincide with years with a great number of good performers (see Figure 6.1), and years when performance improved (see Table 6.3). The years with the lowest innovation score (2002, 2007, and 2008) coincide with some of the years that occurred a performance decline in the industry, as previously mentioned in section 6.4.1 (see Table 6.3). Concerning the years with the highest innovation score (1997 and 2004), it can be concluded that a good innovation score for a few companies does not necessarily coincide with years of particularly good performance for the whole industry.

Analyzing the performance pattern of innovator companies over time, we can conclude that high levels of performance typically precede the innovative status of companies. In particular, the companies that were innovative more than once during the period studied (13) were able to keep high levels of performance in most years for which data was available.

Next, we explored the factors that promote innovation. We used a logistic regression, where the dependent binomial variable represented the company status: innovators (1) or followers (0). The regressors were company size, headquarter location, R&D engagement, as well as the average value of profitability and value added. Table 6.8 reports the estimates from the logistic model, the coefficients, standard errors, Wald statistics, and p-values. Note that the total number of observations included in the model is 97 corresponding to all companies analyzed.

Table 6.8: Logistic regression estimates

Variable	Coef.	Std. error	Wald Test	p-value
Size	-0.158	0.364	0.188	0.665
Lisbon area	0.434	0.545	0.635	0.426
R&D engagement	-0.294	0.574	0.262	0.609
Profitability	0.297	0.114	6.749	0.009
Value added	0.039	0.026	2.138	0.144
Constant	-3.378	1.071	9.953	0.002

Table 6.8 shows that company size, R&D engagement, and location are not statistically significant at a 5% level. Concerning the financial status proxied by the profitability variable, the positive and significant coefficient confirms that it is critical to promote innovation. Note that the coefficient for a variable in a logistic regression represents the change in the odds loga-

rithm. Therefore, for the profitability indicator, a unit increase in the value of this variable increases the probability of being innovator by the factor 1.35 ($= e^{0.297}$). In terms of labor productivity proxied by the value added variable, no significant relationship with innovation could be identified. This means that it is not a critical factor to explain innovation although the positive coefficient may suggest that it can occasionally contribute to enhance innovation.

The overall model is statistically significant (χ^2 test with p-value of 0.000), with a Nagelkerke R^2 equal to 0.23. The overall percentage of observations classified correctly is 79.4% $\left(\frac{5+72}{97}\right)$. This value is larger that the proportional chance criterion of 64.9% $\left((\frac{22}{97})^2 + (\frac{75}{97})^2\right)$, which shows that the logistic regression is a good option to model the factors that promote innovation.

6.5 Summary and conclusions

The purpose of this chapter was to develop a quantitative approach to evaluate the financial soundness of construction companies, and to identify the drivers that promote performance improvements and innovation. To assess contractors performance we used a composite indicator calculated using DEA and complemented with bootstrapping, to obtain a robust estimate of performance. The study revealed that the performance of Portuguese construction companies increased considerably during 1996-2000, which coincides with the prosperous period of the Portuguese CI. From 2001 to 2009 the performance of construction companies slightly increased. This period corresponds to the slowing down of the CI in Portugal, and include the years when performance declined (2001, 2004, 2007, 2008).

To investigate the impact of the national economic context, company size, location and engagement on R&D projects on performance we used a panel data truncated regression. Regarding the factors considered more influential, we concluded that company performance is significantly affected by the national economic context, and that small specialized companies and large companies tend to achieve the best performance levels.

Concerning the assessment of innovation, this chapter proposed an enhanced approach to identify innovative companies, to quantify the extent of innovation and to explore the factors that drive innovation. It was concluded that keeping the innovative status for long periods is difficult, although a few companies were able to keep their innovative status in consecutive years, maintaining a good performance level over the years. Concerning the factors that promote innovation, the results of the assessment suggested that the innovators are typically companies with high levels of profitability.

This study aimed to provide insights that may help managers, or central and local government planners, to define strategies for performance improvements of construction companies. It is essential to adopt effective strategies to improve the competitiveness of the CI in a global construction market. Although a financial diagnosis is of particular interest to organizations, it should be regarded as a first step in a company-wide assessment. The methodology developed in this chapter can be easily applied to other areas that are also important for CI companies, such as marketing performance or operational performance. Finally, an interesting topic for further research concerns a detailed analysis of the innovative companies to identify the management practices that may promote performance improvements.

Chapter 7

Performance trends in the Construction Industry worldwide: an overview of the turn of the century

7.1 Introduction

Construction is one of the largest sectors in terms of Gross Domestic Product (GDP) and employment in the world and highly contributes to economic development. As explained by Crosthwaite (2000), construction industry (CI) produces the facilities that accommodate a wide variety of human activities, and the infrastructure that connects these facilities into an increasingly complex network. The facilities are needed for the production of all other goods and services, starting from those needed by other producers and ending with those needed by the ultimate consumers.

The CI is both highly competitive and cyclically sensitive (Moscarini and Postel-Vinay, 2009). The CI is no longer a local market, given globalization. Construction companies, mainly from the developed countries, are adopting strategies of internationalization that enable them to benefit from the global market. In particular, some American and European construction companies have moved their entire operations to the Middle East, with lower running costs, more work and opportunities. As pointed out by Ngowi et al. (2005), there are several ways in which construction companies enter the international market, for instance: i) economic booms such as the one resulting from sale of oil, ii) bilateral and multilateral agreements, which set up protocols that enable companies of the participating countries to enter the markets of each other, iii) participation in large international projects, or work for multinational corporations. To reach a competitive position in the globalized construction market, construction companies are increasingly interested in cross-country performance comparisons. International benchmarking is particularly important in the CI, allowing a broader view of the industry. As companies have to continually improve their productivity to remain competitive, they are forced to revise their vision, taking into account the company internal situation, their competitors' strategies, and the evolution of the economic context. Construction companies attempt to implement systematic methods of performance measurement to achieve sustainable growth, profitability and competitive advantage.

In the CI sector the competitive pressures are likely to vary according to construction activity as the different segments of companies serve different economic sectors. The construction of buildings and civil engineering works is undertaken in a similar way worldwide: a general contractor, responsible for delivering the finished project to the owner, subcontracts much of the practical work to specialty trade companies. The specialty trade contractors have particular skills related to all types of construction such as carpentry, painting, plumbing, or electrical work. The building segment includes the general contractors, who build residential buildings, and nonresidential, such as industrial, commercial, and other buildings. The heavy civil engineering contractors build roads, highways, bridges, tunnels, and other projects related to national infrastructure. Consequently, the residential building companies are associated with household demand, the nonresidential and heavy civil companies mainly serve the Government demand, and specialty trade mostly serves industrial sectors. This makes it vital to examine differences in efficiency by construction activity area.

It is also of interest to analyze the hypothesis of convergence in efficiency across regions due to the considerable changes that occurred in the CI worldwide in recent years, and also the spread of managerial methods, including the project management (Walker, 2007). North America and Europe correspond to the largest CI markets, including the world's largest contractors. North America is perhaps the leading region in terms of performance and advanced technology. However, the CI in Asia is booming. Asia offers a huge market for construction and engineering services of all kinds and across all sectors. Demands for housing are constantly growing as the population increases and more people move into cities. All kinds of infrastructure are needed to support these populations and to achieve national development objectives in all the regions. The Asian policies also tend to keep CI busy as governments spend on infrastructure as a way of protecting jobs and boost the economy.

The purpose of this chapter is to develop a methodology to assess the efficiency level of construction companies worldwide, exploring in particular the effect of company geographic location and activity in the efficiency. This chapter also provides insights concerning the convergence in efficiency across regions worldwide. For this purpose, we divided the construction companies into three regions of the world (Europe, Asia and North America), and into the three main construction activities (Buildings, Heavy Civil and Specialty Trade contractors). The methodology adopted involved the estimation of efficiency levels using DEA, complemented with bootstrapping to obtain a robust efficiency estimate. The effect of company location and activity on the efficiency levels was tested using a panel data truncated regression with categorical regressors. The convergence in efficiency across regions was analyzed using the Malmquist index, complemented with bootstrapping, for the estimation of the productivity change of construction companies over time.

The remainder of this chapter is organized as follows. Section 7.2 describes the methods used in this study. Section 7.3 characterizes the sample of CI companies, and presents the performance assessment model. Section 7.4 discusses the results obtained, and the last section concludes.

7.2 Methodology

7.2.1 Efficiency assessment

The methodology adopted in this chapter involves the estimation of worldwide construction companies efficiency with a DEA model under constant returns to scale and with an output orientation, using model (2.9). We estimate the relative efficiency of a construction company in a given year, compared to the best practices observed during the period analyzed. Bootstrapping was used to provide statistical inference regarding the significance of the results obtained. This procedure was implemented using the statistical package R including the FEAR library, developed by Wilson (2008).

To explore the effect of company location and activity on efficiency levels of construction companies a truncated regression with bootstrapping was applied, formulated according to the algorithm #2 proposed by Simar and Wilson (2007). The model uses the bias-corrected efficiency as the dependent variable, and as regressors the categorical variables related to company location and activity. In particular, due to the panel nature of the data we used a panel data truncated model controlling for time effect. The model is specified as follows:

$$E_{jt} = \alpha_o + \eta_t + z_{jt}\beta + \varepsilon_{jt} \tag{7.1}$$

Subscript j represents the j^{th} DMU ($j = 1, \ldots, n$), and subscript t represents the time period ($t = 1, \ldots, l$). η_t is a vector of dummy variables for each year, α_o is the intercept, z_{jt} represents the regressors previously identified, β denotes the regression coefficients and ε_{jt} is the error term with a $N(0, \sigma_\varepsilon^2)$ distribution with a truncation at $(1 - \alpha_o - \eta_t - z_{jt}\beta)$. Note that E_{jt} corresponds to the efficiency of DMU$_j$ estimated using model (2.9) and corrected by bootstrapping.

7.2.2 Productivity change over time

The productivity change of construction companies was estimated using the Malmquist productivity index (MI). As explored by Tulkens and Vanden Eeckaut (1995), the MI can be calculated using three different types of frontiers. The contemporaneous frontier that at time period t includes observations only from that time period, the sequential frontier that at time period t includes all observations up to time period t, and the intertemporal frontier that includes all observations regarding the entire period under analysis. Technological advances in the CI are not likely to be unmade, so we used the sequential frontier that prevents technological regress, conceptually and in terms of measurement. As presented in section 2.3, the Malmquist index, measuring productivity change between time periods t and $t+1$, can be obtained as follows:

$$MI^{t+1,t} = \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} \cdot \left(\frac{D_o^t(x^t, y^t)}{D_o^{t+1}(x^t, y^t)} \cdot \frac{D_o^t(x^{t+1}, y^{t+1})}{D_o^{t+1}(x^{t+1}, y^{t+1})} \right)^{\frac{1}{2}} = EC \cdot TC \tag{7.2}$$

In a sequential assessment, $D_o^t(x^t, y^t)$ corresponds to the output distance function estimated for a DMU in period t, denoted by (x^t, y^t), relative to the technology up to period t, denoted by D_o^t. The distance functions embodied in the MI can obtained as the inverse of the optimal solution to model (2.9), i.e. $D_o^t(x^t, y^t) = \frac{1}{h_{jo}^*}$, as explained by Fare and Lovell (1998). As mentioned in section 2.3, the first component (EC), outside the square bracket, reflects the relative change in efficiency between periods t and $t+1$. The second component (TC), corresponding to the geometric mean of the two ratios in square brackets, reflects the relative distance between the frontiers of the two periods.

The values of $MI^{t+1,t}$ may be greater, equal or smaller than one, depending on whether productivity growth, stagnation or decline occurred between periods t and $t+1$. A similar interpretation can be applied to EC. In terms of the TC component, it may be greater or equal than one, which corresponds to technological progress, or stagnation. In a sequential assessment, the most efficient companies in a given year may not reach the highest productivity levels observed in previous years, such that the sequential frontier will only be defined from DMUs of previous time periods. In these circumstances, the performance decline in that year will be captured by the EC component, meaning that the best practice standards of that year are below of what was observed before.

To evaluate the robustness of the estimates obtained for each company we constructed confidence intervals for the MI using bootstrapping. In particular, we used the bootstrapping method proposed by Simar and Wilson (1999) specially designed for the computation of the MI.

7.3 The data

The data used in the study came from the OSIRIS database, a Bureau Van Dijk database which provides financial data on listed companies around the world for a period of more than 20 years (for further details, www.bvdep.com).

We analyzed a sample of 118 companies over the period 1995 to 2003. To construct a balanced panel data, we studied only the companies with data for all these years. The 118 companies analyzed belong to 18 countries from three continents (Europe, Asia and North America) and operate in the three main CI activity profiles (Building construction, Heavy & Civil engineering construction, Specialty Trade contractors). Note that in our sample, building companies mainly represent the residential building segment (56 companies) rather than the nonresidential construction (9 companies). This classification is based on the North American Industry Classification System at a 3-digit level. Table 7.1 details the geographic location and the activity profile of the companies included in the sample.

In this chapter, we used a performance assessment model that characterizes contractors from a financial perspective. The variables selected for a financial benchmarking model should capture the multidimensional nature of the CI activity, and the overall company performance. Based on the review of CI literature, it can be concluded that some of the most critical financial ratios to measure contractor performance are: liquidity, leverage, and profitability. For instance, to measure the three critical financial performance ratios, Altman (1968) proposed working capital, retained earning, and sales, respectively. Kangari et al. (1992) proposed revenue to working capital, return on total assets, and total assets to sales, respectively. Pilateris and McCabe (2003) proposed accounts receivable and working capital to measure liquidity, fixed assets, total debts, and net worth for leverage, and net income, and operating profit for profitability. Beyond the three critical financial performance ratios, Severson et al. (1994) proposed another financial ratio related to cost accounting, measured by underbillings. To measure liquidity, leverage, and profitability the authors used current liabilities, retained earnings, and net profit, respectively.

Table 7.1: Sample characteristics

Location	**Activity Profile**			**Total**
	Buildings	**Heavy&Civil**	**Trade**	
Belgium	0	1	0	1
Denmark	2	0	0	2
France	2	4	0	6
Germany	1	3	0	4
Greece	0	1	0	1
Italy	0	3	0	3
Netherland	1	1	0	2
Norway	0	1	0	1
Portugal	1	3	0	4
Spain	3	2	0	5
Sweden	3	0	0	3
UK	13	3	0	16
Japan	4	7	3	14
Korea, Rep.	26	2	4	32
Philippine	2	0	0	2
Thailand	2	1	0	3
Canada	1	1	1	3
USA	4	7	5	16
Total in Europe	**26**	**22**	**0**	**48**
Total in Asia	**34**	**10**	**7**	**51**
Total in N. America	**5**	**8**	**6**	**19**
Total	**65**	**40**	**13**	**118**

The performance assessment model specified in this chapter includes financial variables representing liquidity, leverage, profitability, and cost accounting. The model uses three inputs and one output. The inputs are: total current liabilities (TCL) to measure liquidity, shareholders' funds (SF) to proxy leverage, and cost of goods sold (CGS) to proxy cost accounting. The output is the net value of sales, which is a profitability measure. The choice of these variables was constrained by data availability.

In terms of variables definition, CGS involves all costs directly allocated to production, including material consumption, wages and salaries relating to the production process, as well as other related production expenses such as rents. SF includes the total share capital, profits retained and reserves. TCL represents the company debts or obligations that are due within one year and includes short term debt, accounts payable, accrued liabilities and other debts. Net sales correspond to the amount of sales generated by a company after the deduction of returns, allowances for damaged or missing goods and any discounts allowed. All the variables were measured in million US dollars. We used nominal values, not adjusted by purchase power parity, because the DEA model has monetary values in all variables, and assumes CRS. Therefore, there is no need to deflate and adjust for purchase power parity, as all input and output variables are equally affected.

Only the companies with non-negative shareholders' funds and net sales were analyzed. Table 7.2 reports the mean and the standard deviation of the variables for the years analyzed.

Table 7.2: Descriptive statistics of the data

Year	**SF (mil$)**		**TCL (mil$)**		**CGS (mil$)**		**Sales (mil$)**	
	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
1995	113.0	667.8	144.7	588.0	68.1	163.0	164.4	465.0
1996	101.5	617.0	134.8	557.4	79.6	184.4	151.1	411.1
1997	95.0	611.6	106.5	447.1	66.8	171.6	129.7	370.9
1998	107.4	677.2	111.0	435.7	67.4	161.8	130.5	365.5
1999	113.7	661.0	124.9	473.7	74.4	175.1	147.5	417.1
2000	103.0	618.0	107.1	409.1	65.4	154.1	130.8	373.8
2001	86.6	472.2	92.2	354.2	61.4	135.0	196.0	953.9
2002	96.9	538.2	94.3	371.4	72.5	156.4	136.6	415.9
2003	114.1	629.1	98.6	366.0	81.0	171.4	153.7	482.7
All years	103.5	610.2	112.7	444.7	70.7	163.7	148.9	472.9

From Table 7.2 it is possible to observe that the companies analyzed are quite diverse, given the large values of the standard deviation. The shareholders' funds variable exhibits the greatest variation over the years whereas the cost of goods sold is the variable with the smallest variance.

Table 7.3 reports the mean values of the four variables for subsamples with the companies grouped by location and activity. From Table 7.3 we can observe that the largest companies are located in Asia and belong to the heavy civil sector. Conversely, the smallest companies are located in North America and belong to the specialty trade sector.

Table 7.3: Mean values of variables for the companies grouped by location and activity profile

Region/Activity	SF (mil$)	TCL (mil$)	CGS (mil$)	Sales (mil$)
Europe	0.6	1.2	1.7	2.6
North America	0.1	0.1	0.5	0.5
Asia	238.8	258.7	160.8	341.0
Buildings	42.6	51.1	62.9	107.3
Heavy & Civil	228.2	237.1	81.2	236.4
Specialty Trade	24.1	34.8	72.9	83.4

7.4 Results and discussion

7.4.1 Efficiency assessment

The first stage of the assessment was intended to assess the efficiency level of construction companies worldwide. The efficiency score for each company in each year was estimated based on a comparison with a pooled frontier representing the best practices observed in the 9 years analyzed. This approach is meaningful as construction companies operate across regions worldwide. We further estimate efficiency considering a DEA model with CRS as this provides a measure of overall technical efficiency and we are interested in this rather than in the components (which variable returns to scale would provide) namely pure technical and scale efficiency. Table 7.4 reports the average of the original CRS efficiency scores as well as the average of the bias-corrected efficiency scores, the standard deviations, and the 95% confidence intervals of each company in each year analyzed.

Table 7.4: Results of original and bootstrapped efficiency estimates

Year	Original Eff	Bias-corr Eff	St. dev	Lower bound	Upper bound
1995	0.492	0.461	0.089	0.436	0.485
1996	0.488	0.460	0.081	0.435	0.482
1997	0.503	0.474	0.082	0.449	0.496
1998	0.473	0.451	0.088	0.428	0.469
1999	0.473	0.446	0.102	0.423	0.468
2000	0.460	0.435	0.099	0.412	0.455
2001	0.468	0.444	0.087	0.421	0.464
2002	0.473	0.448	0.089	0.425	0.468
2003	0.475	0.448	0.097	0.425	0.469
All years	0.478	0.452	0.090	0.428	0.473

From Table 7.4 we can verify that the results point to a low efficiency level during the period analyzed. It is also important to note that the efficiency levels remained relatively stable over the years. In particular, from 1995 to 2003 the results indicate a performance decline of approximately 3%. These results are supported by both original efficiency scores and bias-corrected efficiency scores. The difference between the two estimates is, on average, 2.6%, which is a relatively small difference.

From Table 7.4 we can also verify that the bias-corrected efficiency estimate is within relatively narrow confidence intervals for all the DMUs, i.e. the lower and upper bounds of the intervals are relatively close. This provides statistical confidence for the bias-corrected efficiency. It is also possible to observe that the original efficiency score (not corrected for bias) lies close to the upper bound of the confidence intervals for all the DMUs. This indicates that the original efficiency score over-estimates the true efficiency. We further observed that the bias estimates are for each company larger that the standard error estimates. The bias-corrected estimates were preferred to the original efficiencies since they represent a more accurate estimation of the true efficiency. These are the values used in the remainder of this study.

Next, we study some factors that potentially explain the spread in the efficiency levels observed during the period analyzed. In particular, we analyze the effect of company location and activity on the efficiency levels. We first plotted in Figure 7.1 the bias-corrected efficiency scores obtained for the 118 companies in the 9 years analyzed, exhibited by the combination of location and activity. Note that Europe only has 2 of the 3 construction activities, namely building and heavy civil companies. Figure 7.1 also plots the 95% confidence intervals of the bias-corrected efficiency for each company, as well as the average of the bias-corrected efficiency for each group (signed with a circle). The companies in each group are ordered by the bias-corrected efficiency.

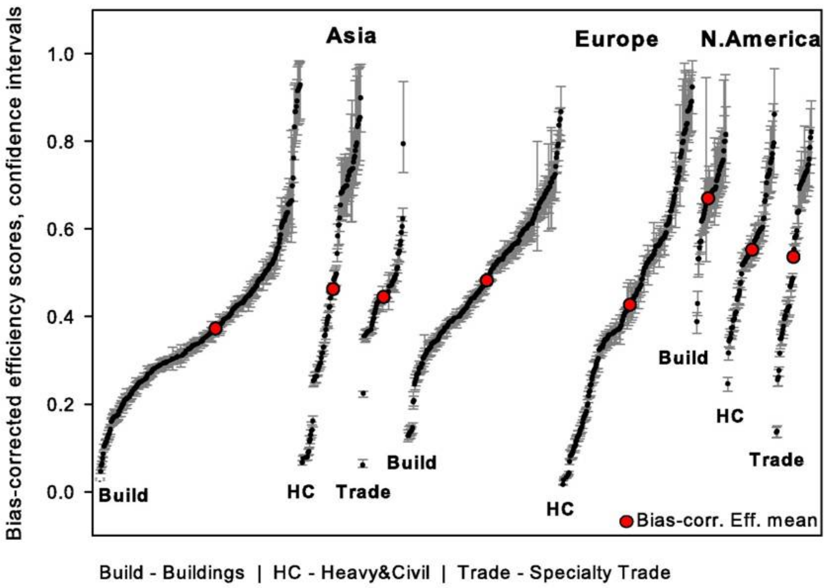

Figure 7.1: Bootstrapping results of the bias-corrected efficiencies by region and activity

From Figure 7.1, we can verify noticeable differences in efficiency levels of construction companies from different regions of the world. In particular, North American companies seem to have the best performance, on average, as well as the smallest efficiency spread in the industry. Concerning the different construction activities, in North America and Europe building companies seem to performance better than the other companies, but the converse occurs in Asia, with speciality trade and heavy and civil companies showing higher efficiency than building companies.

The remainder of this section intends to explore in more detail the effect of region and activity on the efficiency of construction companies. A panel data truncated regression was formulated using the geographic location and activity as regressors, and the bias-corrected efficiency score as the dependent variable. Geographic location and activity profile correspond to categorical variables, with three levels each one. To allow a direct pairwise analysis of regions and to explore interaction effects between activities within regions, the categorical variables had to be coded using an appropriate coding system. The most common approach to deal with categorical variables in regression analysis is to specify dummy variables. Using dummy variables we are able to compare each level of a variable to the reference level. However, we want to compare by pairs the three levels of location and activity variables, which needs a special coding. A possible way of undertaking this is to compare each level to the mean of the subsequent levels of the variable, using a "Helmert" coding. This coding was implemented in the STATA software, developed by StataCorp.

In terms of regions, we compared North America with Asia and Europe, and subsequently compared Europe with Asia. In terms of activities within the regions, in North America and in Asia we compared building companies with heavy civil and specialty trade, as well as heavy civil with specialty trade. In Europe, we compared buildings and heavy civil companies. Table 7.5 reports the results of the panel data truncated model, including the coefficients, standard errors and *p*-values. Note that we also included time dummies to control for year effect. The total number of observations included in the model was 1062 corresponding to the 118 companies in the 9 years analyzed.

The overall regression model is statistically significant (χ^2 test with p-value of 0.000), with a pseudo R-squared equal to 0.146.

Table 7.5: Truncated regression analysis results

Variable	**Coef.**	**Std. error**	**p-value**
North America	0.163	0.014	0.000
Europe	0.021	0.016	0.202
Build in N.America	0.141	0.023	0.000
HC in N.America	0.018	0.031	0.570
Build in Asia	-0.089	0.019	0.000
HC in Asia	0.019	0.032	0.541
Build in Europe	0.060	0.020	0.003
1996	-0.002	0.026	0.943
1997	0.014	0.027	0.604
1998	-0.011	0.026	0.660
1999	-0.017	0.027	0.535
2000	-0.029	0.026	0.260
2001	-0.019	0.025	0.448
2002	-0.014	0.026	0.584
2003	-0.014	0.026	0.576
Constant	0.486	0.007	0.000

Firstly, analyzing the year dummies results in Table 7.5 we confirm that no significant differences in efficiency of construction companies occurred over the years when compared with the reference year of 1995.

The results in Table 7.5 also confirm the results previously discussed. A positive and significant coefficient for North America was obtained, which means that, on average, North American companies are the most efficient when compared with the other companies in the sample. Although the coefficient for Europe is positive, it is not significant, meaning that there is no significant difference in efficiency between European and Asian companies. Note that some relevant CI players in Asia, namely China, India, and Middle East countries, such as United Arabs Emirates are not included in our sample, which may influence the low efficiency scores observed in Asia. Recently, these countries have witnessed a rapid expansion of construction activity, becoming players of international importance in the CI.

Analyzing the interaction effects of activities within regions, we concluded that in North America the building companies are significantly more efficient than heavy civil and specialty trade segments. Between heavy civil and specialty trade companies there is no significant difference. In Europe, building companies perform better than heavy civil. In Asia, building companies are the worst performing companies when compared with the other activities, as we found a negative and significant coefficient for building companies in Asia. Comparing the efficiency between heavy civil and specialty trade companies in Asia there is no significant difference.

The CI comprises works primarily engaged in the construction of buildings and engineering projects. Building and heavy civil companies are responsible for the entire construction projects, although heavy and civil projects are usually larger in scope than building projects. The specialty trade contractors have no responsibility for the project as a whole, in particular they obtain orders from general contractors, or owners. In residential building segment (the building segment most represented in this study) the clients are from the private sector, i.e. the homebuyers, whereas in heavy civil engineering the clients mainly correspond to government entities. As there is a large supply of homes on the market, either in location or design for example, it means that clients have available a wide choice, and as such do not depend on specific building companies to be able to make their purchase. Buying a house is usually a very thoughtful act, as it constitutes one of life's major purchases, which leads clients to be very demanding. In turn, building companies must be able to meet the requirements and aspirations of clients to achieve competitive advantage over their competitors. Furthermore, the level of regulation in the residential market is rigorous in most countries since poor quality work can be costly to owners and potentially hazardous. The competitive environment affecting the residential building segment may explain the greater similarity among companies, and the higher efficiency levels observed both in Europe and North America.

Another aspect that may influence the efficiency level of the building companies in Europe and North America is that homebuilding is a relatively stagnant market. This means that buildings have to face more adverse conditions to survive in the market, which encourages excellence. In Asia, the homebuilding segment exhibited a solid growth in recent years, implying the existence of ample business opportunities. It is interesting to note that the low efficiency levels observed in Asian building companies are mainly associated with companies from Thailand and Philippines, where there is room to improve in terms of construction practices. These aspects may explain the larger dispersion among building companies in Asia and a lower efficiency level, on average.

7.4.2 Productivity change over time

This section of the study explores the hypothesis of convergence in efficiency levels across regions. For this purpose, we estimated the productivity change of construction companies from different regions of the world over the period analyzed. The assessment was done considering each region separately, and a sequential frontier. Table 7.6 summarizes the mean of the Malmquist index for the companies grouped by location over the years.

Table 7.6: Malmquist index, efficiency change and technological change by region

Period	**Malmquist Index**			**Efficiency Change**			**Technological Change**		
	Europe	Asia	N. Amer.	Europe	Asia	N. Amer.	Europe	Asia	N. Amer.
95/96	1.04	1.04	1.03	0.99	0.97	1.01	1.05	1.07	1.02
96/97	1.02	1.05	0.99	0.99	1.04	0.98	1.03	1.00	1.02
97/98	1.00	0.92	0.99	0.99	0.92	0.99	1.00	1.00	1.00
98/99	0.98	1.04	1.00	0.96	1.02	1.00	1.02	1.01	1.00
99/00	1.02	1.02	1.00	0.99	1.01	0.99	1.03	1.01	1.00
00/01	1.00	1.11	0.99	0.99	1.11	0.99	1.02	1.00	1.00
01/02	1.05	1.04	1.02	1.05	1.04	1.01	1.00	1.00	1.01
02/03	1.02	1.05	0.98	1.02	1.03	0.97	1.00	1.02	1.01
Mean	1.01	1.03	1.00	1.00	1.02	0.99	1.02	1.02	1.01

To test the robustness of the MI results for individual companies we used bootstrapping. Table 7.7 displays for each location the number of companies that increased productivity (both bounds of the confidence intervals of MI are higher than one), decreased productivity (both bounds of the confidence intervals less than one), and maintained productivity (confidence intervals include the value of one).

Table 7.7: Number of companies that increased, decreased or maintained productivity by region

Period	**Europe**			**Asia**			**North America**		
	#inc	#equal	#dec	#inc	#equal	#dec	#inc	#equal	#dec
95/96	23	7	18	25	12	14	8	9	2
96/97	20	10	18	21	16	14	6	4	9
97/98	24	6	18	12	13	26	6	9	4
98/99	21	8	19	17	17	17	7	5	7
99/00	20	9	19	20	15	16	7	5	7
00/01	20	5	23	31	12	8	5	3	11
01/02	20	10	18	21	9	21	10	6	3
02/03	20	12	16	21	11	19	1	7	11

Analyzing Table 7.6 we can verify that Asia is the region that had the largest productivity improvement during the period analyzed. In particular, Asia improved productivity in all years, with the exception of 1998. This decline is confirmed by the large number of companies (26 out of 51) that decreased productivity (see Table 7.7). It is interesting to note that this decline in productivity occurred immediately after the Asian stock market collapse that triggered the Asian financial crisis. The productivity decline was due to a considerable efficiency decrease (efficiency change index equal to 0.92). In particular, we observe that the best practice companies also declined efficiency between 1997 and 1998, such that the sequential frontier of 1998 is only constituted by companies from previous years.

In Europe productivity improved, on average, less than in Asia, and in 1999 productivity declined. It is interesting to note that this period coincides with the slow down of European economic activity due to a decrease in exports that began in 1998 following the Asian financial crisis. As a consequence, most companies decreased their efficiency levels (efficiency change index equal to 0.96) although the sequential frontier moved towards more productive levels in 1999 (technological change index equal to 1.02). We can conclude that most companies were adversely affected by the slow down of the economic activity, although a few, mostly from the United Kingdom, were able to increase their productivity and expand the best-practice frontier, despite the adverse economic context.

In North America, the productivity of construction companies remained stable over the years. The most significant productivity decline occurred in 2003, with the majority of companies exhibiting a significant productivity decline (11 out of 19). This decline may be a consequence of the minor recession that affected North America in 2001. In this year, the economic activities slowed down, causing a residential disinvestment.

The productivity results confirm the hypothesis of convergence in efficiency levels across regions. In particular, North America is the region with the highest level of efficiency, as confirmed by the regression results, and with stable productivity levels over the years. Asia and Europe slightly increased productivity over the years. Nevertheless, we verified that the Asian financial crisis in 1998 had a considerable negative impact on the productivity of Asian construction companies, and also affected European companies one year after. In North American, the crisis in 2001 may cause a decline in the performance of American construction companies. These results provide further insights concerning the general belief that the economic context impacts the performance of the construction sector, as pointed out by, for example, Ngowi et al. (2005).

7.5 Summary and conclusions

This chapter studied the efficiency levels of construction companies worldwide, focusing on the effect of company location and activity in the efficiency as we were interested in exploring if different operating environments that exist in different regions (e.g. regulation, unions, etc) have a significant impact on the efficiency of construction companies. Another important contribution of this chapter was the evaluation of convergence in efficiency across regions. This study is to the best of our knowledge the first to undertake international benchmarking comparisons in the CI sector. This is of particular interest to support the design of company policies in an increasingly global and competitive construction market. Construction companies have to gain a deep understanding of the evolving environment, and to shape their strategies in order to achieve a competitive position.

The efficiency levels of construction companies worldwide were explored using DEA, complemented with bootstrapping, to obtain a robust efficiency estimate. We observed that the efficiency levels were particularly low and kept stable during the period analyzed. Then, we explored the relationship between the efficiency levels and company location and activity, using a panel data truncated regression with categorical factors: the geographic location and activity profile. We used a "Helmert" coding in the specification of the regression to enable a detailed analysis of these effects. The results indicated that the efficiency of North American companies is higher than the European and Asian counterparts. We also concluded that in North America and Europe the building companies have higher efficiency levels than the other companies, but the converse occurs in Asia, where building companies performance is worse than in heavy civil and trade companies. The greater efficiency spread of building companies in Asia may be due to the recent growth of the market in this region of the world, such that ample business opportunities enable less efficient companies to remain viable and operate in the construction sector.

Concerning convergence in efficiency levels, we concluded that the North American companies have the highest efficiency levels and their productivity remained stable over the years, whereas the productivity of Asian and European companies improved slightly over the years. This points to a convergence in efficiency levels across regions. Finally, the results of this study confirm the existence of a relationship between the economic context and the performance of construction sector. In particular, we observed that regional economic crisis have a negative impact on the productivity of construction companies.

Chapter 8

Conclusions and directions for future research

8.1 Summary and conclusions

The main purpose of this thesis was to develop a robust approach for performance assessment and improvement in construction companies to enable facing the highly competitive environment in the CI. The first two chapters (4 and 5) of the empirical part of the thesis developed models to evaluate performance at a company level, providing insights concerning the strengths, weaknesses and targets for improvement. The two subsequent chapters (6 and 7) explored performance trends at industry level, identifying the factors that promote performance improvement. The aims, models and methodologies developed, and the major conclusions drawn from the empirical part of this thesis can be summarized as follows.

In chapter 4, it was proposed a performance assessment system based on several performance indicators covering financial, operational and proposal attributes. The system functions as a decision support system to facilitate the selection of the best company to be contracted among competitive bids within e-marketplaces. As the system enables a broad view of the company performance, it can also be used as a tool to guide construction companies to identify weaknesses, strengths, and areas of potential improvement.

In chapter 5, it was proposed a methodology for performance assessment applying DEA to complement the information provided by a set of key performance indicators available in web benchmarking platforms, which are a tool frequently used by construction companies. The methodology provides managerial insights related to organization overall performance and targets to gradually improve performance for all companies, even for the best-practice companies. This is achieved by the specification of virtual companies, whose activity reflects the decision maker preferences. The methodology proposed is particularity useful for CI organizations involved in benchmarking routines.

To demonstrate the advantages of integrating DEA with KPI benchmark scores, the data available in the icBench platform was used. The results showed that DEA is a powerful tool to complement the KPIs analysis, as DEA evaluates companies allowing them to select their own weights, such that good performance in specific areas is valued, and differentiation strategies rewarded. Therefore, DEA enables obtaining rankings more aligned with company strategy. In addition, the DEA analysis enables comparing the inefficient companies with their peers pointing directions for performance improvement. This managerial information is very important to drive company strategy, as the aspects that need more attention can be identified.

In chapter 6, it was proposed a model to evaluate performance within the Portuguese construction industry, identifying the factors that promote performance improvement. In addition, it was developed an enhanced model to identify innovative companies, to quantify the extent of innovation, and to explore the factors that drive innovation. These new models are important to define sustained strategies towards excellence, and to clarify the impact of factors that promote performance improvement in the CI.

The study revealed that after a period of a remarkable performance improvement in the Portuguese CI during the 1990's, the growing trend in performance slowed down in recent years. This result suggests that construction companies should devote greater attention to the implementation of performance improvement practices to cope with the adverse conditions that the Portuguese CI is facing. It was found that company size affects significantly the performance level of construction companies. In particular, small specialized companies and large contractors were considered the best performers. This may indicate that the Portuguese construction companies with intermediate scale size should reorganize themselves or redirect strategic options. The application of the model to assess innovation within the CI revealed that it is difficult to maintain an innovative status for long periods, although a few companies were able to keep high levels of performance in consecutive years. In conclusion, Portuguese CI companies need to strengthen their competitive position over time.

In chapter 7, it was assessed efficiency in the CI worldwide, exploring the effect of company location and activity on efficiency. In addition, the hypothesis of convergence in efficiency across regions was analyzed using the Malmquist index. This study is of particular interest to support the design of company competitive policies, aiming to prosper in an increasingly global construction market.

The study was applied to companies from three different regions of the world: Europe, Asia and North America, and belonging to the three main construction activities: buildings, heavy civil and specialty trade contractors. The results revealed that the efficiency of North American companies was higher than the European and Asian counterparts. Another important conclusion points to a convergence in efficiency levels across regions as in North America productivity remained stable, whereas in Asia and Europe productivity improved. In terms of results by activities, it was concluded that in North America and Europe the building companies had higher efficiency levels than the other companies, but the converse occurred in Asia, where building companies' efficiency was worse than in heavy civil and trade companies. This may be due to the recent growth of the Asian homebuilding market, such that ample business opportunities enable less efficient companies to operate in the sector.

8.2 Contributions of the thesis

The performance assessment of construction companies is an issue of vital importance to government planners and managers in order to guide effective strategies to improve the competitiveness of the construction companies in a global market. Despite the considerable amount of research related to performance measurement in the construction industry, there is a lack of more advanced and robust models for benchmarking and dissemination of best practices to cope with today's challenges. This research is a contribution to this field providing innovative models and methodologies for performance assessment and improvement covering both organizational and industry level issues. The major contributions of this thesis are summarized as follows:

- The definition of a set of performance indicators covering various perspectives of company activity: financial, operation and proposal attributes. These indicators are suitable for inclusion in e-marketplaces.
- The design of a performance assessment system, that can help the selection of bids from potential subcontractors.
- The development of a methodology that enables to integrate the DEA

technique with information on KPIs provided by benchmarking platforms.

- The application of a composite indicator model to evaluate the performance of the CI in Portugal over the years, adopting a financial perspective. This model can be easily adapted to evaluate the CI sector of other countries.
- The estimation of an innovation score for construction companies, using an enhanced DEA model to assess innovation within an industry.
- The identification of several factors associated with good performance and innovation in the construction industry.
- The conduction of an international benchmarking comparison of construction companies, covering a time period of one decade.
- The description of the effect of region and activity on the efficiency level of construction companies worldwide.
- The application of the Malmquist Index to provide insights concerning the convergence in efficiency of construction companies across regions.

Overall, this thesis contributes to illustrate how DEA combined with other techniques can be used with respect to a multitude of objectives of performance evaluation and improvement in the construction industry.

8.3 Directions for future research

This section intends to summarize the directions in which future research would be most fruitful.

In the last two chapters of this thesis, the models for assessing performance characterized companies mainly from a financial perspective. Although a financial diagnosis is of particular interest to all organizations, it should be regarded as one of the important issues in a company-wide assessment among others. An interesting topic for future research would be to apply the methodologies developed in other functional areas, such as the assessment of commercial or operational performance, in order to obtain a complete diagnosis of company performance.

This thesis developed methodologies to assess and improve performance of construction industry companies. The construction industry is a sector that plays a fundamental role in the economic development of countries, and supports many other economic activities. For future research, it would be interesting to extend the performance assessment to other sectors that may be associated with the construction industry (e.g. banking, tourism or real-estate), as well as to investigate the relationship between the performance of the construction industry and the performance changes in these sectors over time. This would allow to define enhanced strategies for performance improvement, and to disseminate the best practices observed in other sectors.

This thesis predominantly used quantitative approaches to evaluate company performance. It would be interesting to complement this research with a detailed qualitative analysis of the benchmark or innovative companies in order to identify managerial and operational practices that promote good levels of performance. It would also be important to promote a collaborative environment in the sector that would benefit all companies and enable spreading best practices in the industry.

This thesis developed methodologies for performance assessment and improvement based on the Data Envelopment Analysis technique. It would be interesting to compare the results obtained in this thesis with those from alternative frontier methods, such as Stochastic Frontier Analysis. Further analysis using other frontier methods could provide additional validation of the results and increase managerial confidence.

References

Abudayyeh, O., Zidan, S. J., Yehia, S., and Randolph, D. (2007). Hybrid prequalification-based, innovative contracting model using AHP. *Journal of Management in Engineering*, 23(2):88–96.

AEP (2007). Construction industry. Conjuncture report, Chamber of Commerce & Industry (AEP), Porto, Portugal.

Aigner, D., Lovell, C., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production models. *Journal of Econometrics*, 6(1):21–37.

Akinsola, A. O., Potts, K. F., Ndekugri, I., and Harris, F. (1997). Identification and evaluation of factors influencing variations on building projects. *International Journal of Project Management*, 15(4):263–267.

Alarcon, L. F. and Mourgues, C. (2002). Performance modeling for contractor selection. *Journal of Management in Engineering*, 18(2):52–60.

Albino, V. and Garavelli, A. C. (1998). A neural network application to subcontractor rating in construction firms. *International Journal of Project Management*, 16(1):9–14.

Allen, R., Athanassopoulos, A., Dyson, R. G., and Thanassoulis, E. (1997). Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions. *Annals of Operations Research*, 73:13–34.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589–609.

Arditi, D. and Chotibhongs, R. (2005). Issues in subcontracting practice. *Journal of Construction Engineering and Management*, 131(8):866–876.

Arditi, D. and Gutierrez, A. (1991). Factors affecting United States contractors performance overseas. *Journal of Construction Engineering and Management*, 117(1):27–46.

Arslan, G., Kivrak, S., Birgonul, M. T., and Dikmen, I. (2008). Improving sub-contractor selection process in construction projects: Web-based subcontractor evaluation system (WEBSES). *Automation in Construction*, 17(4):480–488.

Banker, R. (1993). Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. *Management Science*, 39(10):1265–1275.

Banker, R. (1996). Hypothesis tests using data envelopment analysis. *Journal of Productivity Analysis*, 7(2-3):139–159.

Banker, R., Charnes, A., and Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9):1078–1092.

Barros, C. P. and Peypoch, N. (2009). An evaluation of european airlines' operational performance. *International Journal of Production Economics*, 122(2):525–533.

Bassioni, H. A., Price, A. D., and Hassan, T. M. (2004). Performance measurement in construction. *Journal of Management in Engineering*, 20(2):42–50.

Bassioni, H. A., Price, A. D., and Hassan, T. M. (2005). Building a conceptual framework for measuring business performance in construction: an empirical evaluation. *Construction Management & Economics*, 23(5):495–507.

Beatham, S., Anumba, C., Thorpe, T., and Murray, M. (2002). Utilizing the EFQM excellence model to drive business improvement. In *Proceedings of the 9th International Conference on Concurrent Engineering*, Cranfield University, United Kingdom.

Bubshait, A. and Al-Gobali, K. (1996). Contractor prequalification in Saudi Arabia. *Journal of Management in Engineering*, 12(2):50–54.

Camp, R. C. (1989). *Benchmarking: The search for industry best practices that lead to superior performance.* ASQC Quality Press, 1st edition.

Castro-Lacouture, D., Medaglia, A., and Skibniewski, M. (2007). Supply chain optimization tool for purchasing decisions in B2B construction marketplaces. *Automation in Construction*, 16(5):569–575.

Caves, D., Chirstensen, L., and Diewert, W. (1982). The economic theory of index numbers and the measurement of input, output and productivity. *Econometrica*, 50(6):1393–1414.

Cazals, C., Florens, J., and Simar, L. (2002). Nonparametric frontier estimation: A robust approach. *Journal of Econometrics*, 106(1):1–25.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring efficiency of decision making units. *European Journal of Operational Research*, 2(6):429–444.

Chen, P., Yu, M., Chang, C., and Hsu, S. (2008). Total factor productivity growth in China's agricultural sector. *China Economic Review*, 19(4):580–593.

Cherchye, L., Moesen, W., and Puyenbroeck, T. V. (2004). Legitimately diverse, yet comparable: On synthesizing social inclusion performance in the EU. *Journal of Common Market Studies*, 42(5):919–955.

Cobb, C. and Douglas, P. (1928). A theory of production. *American Economic Review*, Supplement to vol.18:139–165.

Cooper, W. W., Thompson, R., and Thrall, R. (1996). Extensions and new developments in DEA. *Annals of Operations Research*, 66:3–45.

Costa, D. B., Formoso, C. T., Kagioglou, M., Alarcon, L. F., and Caldas, C. H. (2006). Benchmarking initiatives in the construction industry: Lessons learned and improvement opportunities. *Journal of Management in Engineering*, 22(4):158–167.

Costa, J. M. and Horta, I. M. (2007). *Indicadores de Desempenho e Produtividade da Indústria da Construção - Resultados 2005*. Porto, Portugal, FEUP/IMOPPI edition.

Costa, J. M., Horta, I. M., Guimarães, N., Nóvoa, H., Cunha, J. F., and Sousa, R. (2007). icBench - a benchmarking tool for portuguese construction industry companies. *International Journal for Housing Science and Its Applications*, 31(1):33–41.

Crosthwaite, D. (2000). The global construction market: A cross-sectional analysis. *Construction Management and Economics*, 18(5):619–627.

CSIR, B. E. U. (2006). *An Architect Guide to Designing for Sustainability.* Commonwealth Association of Architects Developmental Study. Pretoria, South Africa.

Curtin, J. J. (1993). *The basic bond book.* Associated General Contractors of America and National Association of Surety Bond Producers, Washington, D.C.

Darvish, M., Yasaei, M., and Saeedi, A. (2009). Application of the graph theory and matrix methods to contractor ranking. *International Journal of Project Management*, 27(6):610–619.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, 19(3):273–292.

Dyson, R., Allen, R., Camanho, A., Podinovski, V., Sarrico, C., and Shale, E. (2001). Pitfalls and protocols in DEA. *European journal of operational research*, 132(2):245–259.

Dyson, R. G. and Thanassoulis, E. (1988). Reducing weight flexibility in data envelopment analysis. *Journal of Operational Research Society*, 39(6):563–576.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26.

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

Egan, J. (1998). Rethinking construction. Technical report, Construction Task Force. Dept. of the Environment, Transport and the Regions, London, United Kingdom.

El-Mashaleh, M. (2010). Decision to bid or not to bid: A data envelopment analysis approach. *Canadian Journal of Civil Engineering*, 37(1):37–44.

El-Mashaleh, M., Minchin, R., and O'Brien, W. (2007). Management of construction firm performance using benchmarking. *Journal of Management in Engineering*, 23(1):10–17.

El-Mashaleh, M., O'Brien, W. J., and Minchin, R. (2006). Firm performance and information technology utilization in the construction industry. *Journal of Construction Engineering and Management*, 132(5):499–507.

El-Sawalhi, N., Eaton, D., and Rustom, R. (2007). Contractor prequalification model: State-of-the-art. *International Journal of Project Management*, 25(5):465–474.

Elazouni, A. M. and Metwally, F. G. (2000). D-SUB: decision support system for subcontracting construction works. *Journal of Construction Engineering and Management*, 126(3):191–200.

Fare, R., Grosskopf, S., Lindgren, B., and Roos, P. (1994a). Productivity developments in swedish hospitals: a malmquist output index approach. In *Charnes, A.; Cooper, W.W.; Lewin, A. and Seiford, L. (Eds.) Data Envelopment Analysis: Theory Methodology and Applications*, pages 253–272. Kluwer Academic Publishers, Boston.

Fare, R., Grosskopf, S., Norris, M., and Zhang, Z. (1994b). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review*, 87(5):1040–1043.

Fare, R. and Lovell, C. (1998). Measuring the technical efficiency of production. *Journal of Economic Theory*, 19(1):150162.

Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, pages 252–267.

Fiegenbaum, A. and Karnani, A. (1991). Output flexibility-A competitive advantage for small firms. *Strategic Management Journal*, (12):101–114.

Fried, H. O., Lovell, C. A. K., and Schmidt, S. S. (2008). *The measurement of productive efficiency and productivity growth*. Oxford University Press, New York, United States.

Fung, M., Wan, K., Hui, Y., and Law, J. (2008). Productivity changes in chinese airports 1995-2004. *Transportation Research Part E*, 44(3):521–542.

Halkos, G. E. and Tzeremes, N. G. (2007). Productivity efficiency and firm size: An empirical analysis of foreign owned companies. *International Business Review*, 16(6):713–731.

Haney, A. B. and Pollitt, M. G. (2009). Efficiency analysis of energy networks: An international survey of regulators. *Energy Policy*, 37(12):5814–5830.

Hartmann, A., Ling, F., and Tan, J. (2009). Relative importance of subcontractor selection criteria: Evidence from Singapore. *Journal of Construction Engineering and Management*, 135(9):826–832.

Hatush, Z. and Skitmore, M. (1998). Contractor selection using multicriteria utility theory: An additive model. *Building and Environment*, 33(2-3):105–115.

Hayami, Y. and Ruttan, V. (1971). *Agricultural Development: An International Perspective*. Johns Hopkins University Press, Baltimore, United Sates.

Holt, G. D. (1996). Applying cluster analysis to construction contractor classification. *Building and Environment*, 31(6):557–568.

Johnes, J., Izzeldin, M., and Pappas, V. (2009). The efficiency of islamic and conventional banks in the Gulf Cooperation Council (GCC) countries: An analysis using financial ratios and data envelopment analysis. Working paper No. 006069, Lancaster University Management School, Economics Department.

Kagioglou, M., Cooper, R., and Aouad, G. (2001). Performance management in construction: A conceptual framework. *Construction Management & Economics*, 19(1):85–95.

Kale, S. and Arditi, D. (1998). Business failures: Liabilities of newness, adolescence, and smallness. *Journal of Construction Engineering & Management*, 124(6):458.

Kangari, R., Farid, F., and Elgharib, H. M. (1992). Financial performance analysis for construction industry. *Journal of Construction Engineering and Management*, 118(2):349–361.

Kaplan, R. and Norton, D. (1992). The Balanced Scorecard - measures that drive performance. *Harvard Business Review*, 70(1):71–79.

Ko, C., Cheng, M., and Wu, T. (2007). Evaluating sub-contractors performance using EFNIM. *Automation in Construction*, 16(4):525–530.

Koopmans, T. (1951). *Activity analysis of production and allocation*. John Wiley and Sons, New York, United States.

Korde, T. (2005). State-of-the-art review of construction performance models and factors. In *Construction Research Congress 2005: Broadening Perspectives*, San Diego, California.

Kumaraswamy, M. M. and Matthews, J. D. (2000). Improved subcontractor selection employing partnering principles. *Journal of Management in Engineering*, 16(3):47–57.

Kuosmanen, T. (2002). Modeling blank data entries in data envelopment analysis. EconWPA Working paper in econometrics No. 0210001.

Latham, M. (1994). Constructing the team. Final report, Government / Industry Review of Procurement and Contractual Arrangements in the UK Construction Industry HMSO, London, United Kingdom.

Latruffe, L., Davidova, S., and Balcombe, K. (2008). Productivity change in polish agriculture: An illustration of a bootstrapping procedure applied to Malmquist indices. *Post-Communist Economies*, 20(4):4–49.

Lee, S., Thomas, S. R., and Tucker, R. L. (2005). Web-based benchmarking system for the construction industry. *Journal of Construction Engineering & Management*, 131(7):790–798.

Lovell, C., Pastor, J., and Turner, J. (1995). Measuring macroeconomic performance in the OCDE: A comparison of European and non-European countries. *European Journal of Operational Research*, 87(3):507–518.

Lynch, R. and Cross, K. (1991). *Measure up!* Blackwell, Cambridge.

Ma, J., Evans, D., Fuller, R., and Stewart, D. (2002). Technical efficiency and productivity change of China's iron and steel industry. *International Journal of Production Economics*, 76(3):293–312.

Maes, J., Sels, L., and Roodhooft, F. (2005). Modelling the link between management practices and financial performance. Evidence from small construction companies. *Small Business Economics*, 25(1):17–34.

Maturana, S., Alarcon, L. F., Gazmuri, P., and Vrsalovic, M. (2007). On-site subcontractor evaluation method based on lean principles and partnering practices. *Journal of Management in Engineering*, 23(2):67–74.

McCabe, B., Tran, V., and Ramani, J. (2005). Construction prequalification using data envelopment analysis. *Canadian Journal of Civil Engineering*, 32(1):183–193.

Moscarini, G. and Postel-Vinay, F. (2009). Large employers are more cyclically sensitive. NBER Working paper No. 14740.

Mukherjee, K., Ray, S., and Miller, S. (2001). Productivity growth in large US commercial banks: The initial post-deregulation experience. *Journal of Banking and Finance*, 25:913–939.

Narasimhan, R., Swink, M., and Kim, S. W. (2005). An exploratory study of manufacturing practice and performance interrelationships: Implications for capability progression. *International Journal of Operations & Production Management*, 25(10):1013–1033.

Neely, A., Adams, C., and Kennerley, M. (2002). *The Performance Prism: The Scorecard for Measuring and Managing Business Success*. Prentice Hall, Great Britain.

Ng, S. T. and Skitmore, R. (1999). Client and consultant perspectives of prequalification criteria. *Building and Environment*, 34(5):607–621.

Ngowi, A., Pienaar, E., Talukhaba, A., and Mbachu, J. (2005). The globalisation of the construction industry - a review. *Building and Environment*, 40(1):135–141.

Odeck, J. (2006). Identifying traffic safety best practice: an application of DEA and Malmquist indices. *Omega: The International Journal of Management Science*, 34(1):28–40.

Odeck, J. (2008). How efficient and productive are road toll companies? Evidence from Norway. *Transport Policy*, 15(4):232–241.

OECD (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD publishing edition.

Palaneeswaran, E. and Kumaraswamy, M. (2001). Recent advances and proposed improvements in contractor prequalification methodologies. *Building and Environment*, 36(1):73–87.

Palaneeswaran, E. and Kumaraswamy, M. M. (2000). Contractor selection for Design/Build projects. *Journal of Construction Engineering and Management*, 126(5):331–339.

Pastor, J. M., Prez, F., and Quesada, J. (1997). Efficiency analysis in banking firms: An international comparison. *European Journal of Operational Research*, 98:395–407.

Pilateris, P. and McCabe, B. (2003). Contractor financial evaluation model (CFEM). *Canadian Journal of Civil Engineering*, 30(3):487–499.

Podinovski, V. V. (2001). DEA models for the explicit maximisation of relative efficiency. *European Journal of Operational Research*, 131(3):572–586.

Portela, M. and Camanho, A. S. (2007). Performance assessment of portuguese secondary schools. Working paper in economics No. 07, Faculdade de Economia e Gestão, Universidade Católica Portuguesa.

Ramirez, R. R., Alarcon, L. F. C., and Knights, P. (2004). Benchmarking system for evaluating management practices in the construction industry. *Journal of Management in Engineering*, 20(3):110–117.

Rao, D. S. P., O'Donnell, C. J., and Battese, G. E. (2003). Metafrontier functions for the study of inter-regional productivity differences. Technical Report CEPA Working Papers Series, No. WP012003.

Ray, S. C. (2007). Are some Indian banks too large? An examination of size efficiency in Indian banking. *Journal of Productivity Analysis*, 27(1):41–56.

Russell, J. S. and Skibniewski, M. J. (1990). QUALIFIER-1 contractor prequalification model. *Journal of Computing in Civil Engineering*, 4(1):77–90.

Severson, G., Russell, J. S., and Jaselskis, E. J. (1994). Predicting construction contract surety bond claims using contractor financial data. *Journal of Construction Engineering and Management*, 120(2):405–420.

Sheater, S. J. and Jones, M. C. (1991). A reliable data-based bandwith selection method for kernel density estimation. *Journal of the Royal Statistical Society*, Series B 53(3):683–690.

Shepard, R. (1953). *Cost and production functions.* Princeton University Press, New Jersey, United Sates.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis.* Monographs on Statistics and Applied Probability. Chapman & Hall, London, United Kingdom.

Simar, L. (2003). Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis*, 20(3):391–424.

Simar, L. and Wilson, P. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1):49–61.

Simar, L. and Wilson, P. (1999). Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research*, 115(3):459–471.

Simar, L. and Wilson, P. (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis*, 13(1):49–78.

Simar, L. and Wilson, P. (2002). Non-parametric tests of returns to scale. *European Journal of Operational Research*, 139(1):115–132.

Simar, L. and Wilson, P. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1):31–64.

Singh, D. and Tiong, R. L. K. (2005). A fuzzy decision framework for contractor selection. *Journal of Construction Engineering and Management*, 131(1):62–70.

Soderbom, M. (2004). Size and efficiency in African manufacturing firms: Evidence from firm-level panel data. *Journal of Development Economics*, 73(1):369–394.

Someya, S. (1992). *The role of R&D in construction firms*. Master of sciences in civil engineering, Massachusetts Institute of Technology, USA.

Sonmez, M., Holt, G. D., Yang, J. B., and Graham, G. (2002). Applying evidential reasoning to prequalifying construction contractors. *Journal of Management in Engineering*, 18(3):111–119.

Sousa, R. and Voss, C. (2002). Quality management re-visited: A reflective review and agenda for future research. *Journal of Operations Management*, 20(1):91–109.

Sowlati, T. and Paradi, J. (2004). Establishing the practical frontier in data envelopment analysis. *Omega: The International Journal of Management Science*, 32(4):261–272.

Strigler, G. (1939). Production and distribution in the short run. *Journal of Political Economy*, 47:305–327.

Thiry, B. and Tulkens, H. (1992). Allowing for inefficiency in parametric estimation of production functions for urban transit firms. *Journal of Productivity Analysis*, 3(1-2):45–65.

Thompson, R., Singleton, L., Thrall, R., and Smith, B. (1986). Comparative site evaluations for locating a high-energy physics lan in Texas. *Interfaces*, 16:35–49.

Topcu, Y. I. (2004). A decision model proposal for construction contractor selection in Turkey. *Building and Environment*, 39(4):469–481.

Tserng, H. P. and Lin, P. H. (2002). An accelerated subcontracting and procuring model for construction projects. *Automation in Construction*, 11(1):105–125.

Tulkens, H. and Vanden Eeckaut, P. (1995). Nonparametric efficiency, progress and regress measures for panel data - methodological aspects. *European Journal of Operational Research*, 80(3):474–499.

Walker, A. (2007). *Project Management in Construction*. Blackwell, Oxford.

Wand, M. P. and Jones, M. C. (1995). *Kernel smoothing*. CRC Press.

Wilson, P. (2008). FEAR: A package for frontier efficiency analysis with R. *Socio-Economic Planning Sciences*, 42(4):247–254.

Wong, C., Nicholas, J., and Holt, G. (2003). Using multivariate techniques for developing contractor classification models. *Engineering, Construction and Architectural Management*, 10(2):99–116.

Wong, Y. and Beasley, J. (1990). Restricting weight flexibility in data envelopment analysis. *Journal of Operational Research Society*, 41(9):829–835.

Xue, X., Shen, Q., Wang, Y., and Lu, J. (2008). Measuring the productivity of the construction industry in China by using DEA-based Malmquist productivity indices. *Journal of Construction Engineering and Management*, 134(1):64–71.

You, T. and Zi, H. (2007). The economic crisis and efficiency change: Evidence from the Korean construction industry. *Applied Economics*, 39(14):1833–1842.