Three essays on treatment quality: theory, measures and application in the hospital sector in China

He, Xinju

Date of Award:
2019

HONG KONG BAPTIST UNIVERSITY

Doctor of Philosophy

THESIS ACCEPTANCE

DATE: August 20, 2019

STUDENT'S NAME: HE Xinju

THESIS TITLE: Three Essays on Treatment Quality: Theory, Measures and Application in the Hospital Sector in China

This is to certify that the above student's thesis has been examined by the following panel members and has received full approval for acceptance in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Chairman:
Dr Zhang Fang
Associate Professor, Department of Accountancy and Law, HKBU
(Designated by the Dean of School of Business)

Internal Members:
Prof Chung Kim Sau
Professor, Department of Economics, HKBU
(Designated by the Head of Department of Economics)

Dr Cao Kang Hua
Assistant Professor, Department of Economics, HKBU

External Members:
Prof Kerstens Kristiaan
Professor
Department of Lille Economie & Management
French National Centre for Scientific Research

Dr Zhou Zhongbao
Distinguished Professor
School of Business Administration
Hunan University

Proxy:
Dr Dai Anovia Yifan
Assistant Professor, Department of Economics, HKBU

In-attendance:
Dr Li Sung Ko
Associate Professor, Department of Economics, HKBU

Issued by Graduate School, HKBU

Three Essays on Treatment Quality:

Theory, Measures and Application in the Hospital Sector in China

HE Xinju

A thesis submitted in partial fulfilment of the requirements

for the degree of

Doctor of Philosophy

Principal Supervisor:

Dr LI Sung Ko (Hong Kong Baptist University)

August 2019

DECLARATION

I hereby declare that this thesis represents my own work, which has been done after registration for the degree of PhD at Hong Kong Baptist University, and has not been previously included in a thesis or dissertation submitted to this or any other institution for a degree, diploma or other qualifications.

I have read the university's current research ethics guidelines and accept responsibility for the conduct of the procedures in accordance with the University's Research Ethics Committee (REC). I have attempted to identify all the risks related to this research that may arise in conducting this research, obtained the relevant ethical and/or safety approval (where applicable), and acknowledged my obligations and the rights of the participants.

Signature: He Xinju

Date: August 2019

ABSTRACT

This thesis investigates the treatment quality of medical services in the hospital sector from different angles: theory, measures and application in China. This thesis includes three essays. The first essay (Chapter 2) is a critical review about the quality assessment. It shows that the environmental performance index is suitable for measuring treatment quality. The second essay (Chapter 3) introduces alternative approaches to environmental performance indices to solve the infeasibility problem of current measures. Environmental performance indices are measures to evaluate the production of undesirable outputs relative to desirable outputs. My new measures are more accurate using the sequential frontier and various reference vectors. The last essay (Chapter 4) is an empirical case study in the Chinese hospital sector that examines how the degree of government involvement and the degree of market competition affect the performance of treatment quality. Using the environmental performance indices of Essay 2 to evaluate treatment quality, I find that the performance of treatment quality of Chinese hospitals improved during the 2009-2014 period. Therefore, the marketisation of hospitals and government subsidies contribute to this sustained improvement.

Keywords: *hospital sector, treatment quality, environmental performance index, undesirable outputs, marketisation, government subsidies*

ACKNOWLEDGEMENTS

First, I would like to express my sincere gratitude to my Principal Supervisor, Dr.LI Sung Ko, for his continued support of my PhD study and related research, and for his patience, motivation and immense knowledge. His guidance helped me in my research and my writing of this thesis. I could not have asked for a better advisor and mentor for my PhD study.

Second, I would like to thank the head of my department, my co-supervisor Prof. Cheng Yuk-Shing for his good effort to assist me in both research and other problem in PhD study. With his kindly help, I enjoyed my past several years and learnt a lot from my PhD study.

Third, I would like to thank the rest of my thesis committee: Dr. Cao Kang Hua, Prof. Chung Kim-Sau, Prof. Kristiaan Kerstens, and Prof. Zhou Zhongbao (Their names are listed in alphabetical order) for their insightful comments and encouragement, and for their difficult questions that prompted me to broaden my research from various perspectives.

Finally, I would like to thank my family, especially my wife Runmin and my daughter Runxin for supporting me spiritually throughout the writing of this thesis and my life in general.

Table of Contents

Chapter 1

Introduction

Treatment quality is important in the hospital sector. Compared with other sectors, the hospital sector is special. In most sectors, consumers gain utility directly through consumption. For example, people know that they will be more satisfied by eating delicious food, using a smartphone or travelling. However, medical consumption is closer to human capital investment than consumption (Grossman, 1972). Patients cannot gain utility or even suffer from the treatment process. Patients do not enjoy the treatment process (surgery, medicine) or want redundant medical services. Patients use medical services because they can improve their health status. Thus, patients gain utility not directly from medical consumption, but from the improvement of their health status. This explains why quality in the hospital sector is more important than in other sectors.

However, the method of indicating the quality of medical services is a critical issue in studies of the hospital sector. In this thesis, the quality of medical services refers to “treatment quality”, which is the number of successful treatments of patients relative to the number of deaths during treatments. The measurement of treatment quality is studied from three different angles in this thesis: (i) a critical assessment of existing methods of measuring treatment quality; (ii) proposing new formulas to remedy the weaknesses of a widely adopted index that is suitable for indicating treatment quality; (iii) applying the new formulas in (ii) to provide an answer to a controversy between two strategies of the healthcare reform in China.

Chapter 2 (Essay 1) provides a critical assessment of some measurement methods of treatment quality. From the definition of WHO[1], the quality of care is defined as "the extent to which health care services provided to individuals and patient populations improve desired health outcomes. To achieve this, health care must be safe, effective, timely, efficient, equitable and people-centred." Base on this definition, the quality of care is multidimensional. By pointing out the pros and cons of various measurement methods in the literature of medical services quality, this essay investigates the literature of environmental protection. It is found that the "environmental performance index" (EPI) introduced by Färe et.al. (2004) is both objective and able to handle different dimensions of treatment quality simultaneously. This index is a ratio of a quantity index of economic

---

[1] See detail definition of quality of care from WHO on https://www.who.int/maternal_child_adolescent/topics/quality-of-care/definition/en/

goods and a quantity index of economic bads. Replacing “goods” and “bads” by “successful treatments” and “deaths”, the environmental performance index can be adopted to measure treatment quality.

Chapter 3 (Essay 2) points out a significant weakness of the existing environmental performance index: A solution to computing this index from empirical data is not guaranteed for every observation. This index is well-defined for all observations under very restrictive conditions. Using terms from the field of productivity analysis, the existing formula of this index is output-oriented. This essay explores two directions to solve this infeasibility problem. (i) Observing that the index is a ratio between two output-oriented quantity indices and such indices are proved to be infeasible in some cases, new input-oriented formulas are adopted to construct quantity indices. The environmental performance index computed from these input-oriented quantity indices is well-defined under less restrictive conditions. (ii) Observing that the relationship between “goods” and “bads” is analogous to the productivity between “outputs” and “inputs”, this essay proposes to use any valid productivity index to measure environmental performance. A modified formula of the Malmquist productivity index with “goods” and “bads” is introduced to capture the environmental performance.

Chapter 4 (Essay 3) applies the new formulas in Chapter 3 to examine the effects of government involvement and market competition on the treatment quality of hospitals in China. The techniques of Data Envelopment Analysis (DEA) are adopted to compute the values of these formulas. This essay discusses the market-oriented strategy and government-oriented strategy of healthcare reforms. In academia, there is no consensus on the effectiveness of these two strategies to promote social welfare. The results of this essay show that these two policies are not mutually exclusive and can be adopted simultaneously.

In summary, this thesis contributes to the literature on treatment quality in the following aspects. Chapter 2 points out that the measurement of treatment quality can borrow the formula of the environmental performance index from the field of environmental protection. Chapter 3 introduces two new alternatives that improve the environmental performance index in the current literature. Chapter 4 provides an answer to the controversies between government involvements and market competition. It also justifies the direction of the ongoing reforms in the Chinese hospital sector.

Chapter 2

A Review of Treatment Quality Assessment

This chapter reviews the literature about treatment quality. There are three issues of studies about treatment quality. First, what is treatment quality? People have different views on the definition of treatment quality. Different definitions of treatment quality go different directions of evaluation of quality. Second, how is treatment quality evaluated? When the definition is clear, the methodology of measuring treatment quality based on an appropriate definition is another important issue. Last, how is treatment quality improved? This is an empirical question which is important for policymakers. With appropriate definition and effective measurement, identifying determinants of treatment quality can help improving the performance of hospitals. The rest parts of this chapter are as follows: Section 2.1 discusses the definition of treatment quality. Section 2.2 reviews the methodologies of quality assessment in both production approach and non-production approach. Section 2.3 reviews the studies about treatment quality evaluation in the hospital sector. Section 2.4 describes the output-oriented environmental performance index which can be used to measure treatment quality. Section 2.5 summarizes different ways of modelling the production technology with undesirable outputs. Section 2.6 concludes.

2.1 Definition of Treatment Quality

“Quality” refers to the degree of excellence. In general, there are several dimensions of the quality of a commodity. For example, when better materials are used, or the speed of computation is faster in the consumption of a desktop computer, we say that there is an improvement in quality. Materials and speed of computation are favourable dimensions in the sense that a higher value of a dimension increases the quality of a commodity. In some production process, desirable and undesirable outputs are jointly produced. Quality is reflected by the quantities of desirable outputs relative to the quantities of undesirable outputs. Each desirable output is a favourable dimension. Each undesirable output is an unfavourable dimension. Thus, more desirable outputs and less undesirable outputs is an improvement in the quality of producing the desirable outputs. Treatments of patients in hospitals are of this kind. If the treatment of patients in a hospital can be regarded as utilizing resources to produce desirable outputs (successful treatments) and undesirable outputs (deaths during treatments), the task of assessing quality of the hospital sector is equivalent to evaluating the favourable, and unfavourable dimensions of a production process.

Quality assessment of medical services is a critical issue of the hospital sector. The ultimate target of medical services in the presence of a disease is to cure the patient. *Treatment* is referred to as the process of curing a patient. *Successful treatment* is a full recovery from a disease. The quality of healthcare in the presence of diseases involves anything that influences successful treatments and the utilities of patients during treatments. Hence successful treatments of patients must be included in accessing the quality of healthcare services. When a patient is cured, it is a *desirable outcome*. However, harms to patients, such as injuries and medical errors, during each treatment may exist. When there is a harm to the patient during a treatment, it is an *undesirable outcome*. Since undesirable outcomes affect the health status of a patient adversely, they should be part of the quality assessment of healthcare services. There are other dimensions that affect healthcare services quality, too. The World Health Organization (WHO) gives a board definition of quality in healthcare as "… health care services provided to individuals and patient populations improve desired health outcomes …" which includes six dimensions: 1). Safe: "Delivering health care that minimizes risks and harm to service users, including avoiding preventable injuries and reducing medical errors.” 2). Effective: “Providing services based on scientific knowledge and evidence-based guidelines.” 3). Timely: “Reducing delays in providing and receiving health care.” 4). Efficient: “Delivering health care in a manner that maximizes resource use and avoids waste.” 5). Equitable: “Delivering health care that does not differ in quality according to personal characteristics such as gender, race, ethnicity, geographical location or socioeconomic status.” 6). People-centred: “Providing care that takes into account the preferences and aspirations of individual service users and the culture of their community.”

These six dimensions can be classified into three types of quality: (i) *treatment quality*: The dimension “safe” in treatment is to achieve the desirable outcome and simultaneously minimizing undesirable outcomes. A desirable outcome (i.e., successful treatment) requires accurate diagnosis and correct treatment methods. Thus, the dimension “effective” is related to achieving the desirable outcome. Similarly, injuries and disease usually become more serious when they are not treated appropriately at the beginning stage. Hence “timely” is also crucial for having a desirable outcome. Given a healthcare system, the dimensions “safe”, “effective” and “timely” are related to having more desirable outcomes and less undesirable outcomes during treatments. Thus, the quality related to them is called treatment quality. (ii) *productive quality*: In this thesis, the treatment of patients in a hospital is regarded as utilizing resources to produce desirable outputs and undesirable outputs. The dimension "efficient" describes how outputs are produced from inputs. So, it is called productive quality. (iii) *service quality*: In dimensions "equity" and "people-centred", healthcare services, hospitals and patients are treated as products, firms, and customers. These two dimensions are related to how services are distributed to customers by firms. So, it is called service quality. This thesis focuses on treatment quality. The other two types of quality will be left for future research.

There are priorities for different dimensions. The safety of healthcare is always the most critical dimension for consideration and amenable mortality is the most undesirable in this dimension. Compared with minimizing deaths during treatments, other dimensions are relatively minor. It is not surprising that in some empirical studies, only amenable deaths are considered in evaluating the quality of medical services (Thomas and Hofer, 1999; Hu et.al., 2012; Pitches et.al., 2007). Since the most important dimension for quality of healthcare is "Safety" and amenable deaths are the most undesirable. Rather than dealing with the board definition of quality of healthcare from WHO, this thesis focuses on a narrow concept of treatment quality that refers to improving the favourable dimensions (or desirable outcomes) and avoiding unfavourable dimensions (or undesirable outcomes).

2.2 Assessing Quality with Favourable and Unfavourable Dimensions

Consider the quality of an object in concern can be reflected by some favourable and unfavourable dimensions. There are two cases, one is the non-production approach, and the other one is production approach. This section reviews the frameworks of analysis under these two cases.

2.2.1 Case 1: Non-production Quality Assessment

Suppose there are A favourable dimensions or desirable outcomes related to the quality of an object. Each dimension or outcome can be represented as one good quality indicator g_i, where $i = 1, \dots, M$. The favourable dimensions can be represented by a vector $G = (g_1, \dots, g_M)$. There are L unfavourable dimensions or undesirable outcomes related to treatment quality. Each dimension or outcome can be represented as one bad quality indicator b_j, where $j = 1, \dots, L$. The unfavourable dimensions can be represented by a vector $B = (b_1, \dots, b_L)$. Let the quality of the whole treatment be Q. Then each quality indicator (good or bad) affects Q. By the definitions of good and bad quality indicators, Q is increasing in $g_i,\ i = 1, \cdots, M$, and decreasing in $b_j,\ j = 1, \cdots, L$.

To assess quality with favourable and unfavourable dimensions, one approach is to assign a value to each dimension and assess the quality by examining each dimension separately. However, there is not one single index to represent overall quality performance. To assess the overall quality performance, the examiners must investigate the quality indicators independently and jump to a conclusion by imposing their subjective value judgements. I call this "independent quality assessment" (IQA). To illustrate the method of IQA, suppose there are two objects U_1 and U_2. The i_{th} favourable dimension of U_k can be represented by the vector $G^k = \left(g_1^k, \dots, g_M^k\right),\ k = 1{,}2$. The j_{th} unfavourable dimension of U_k can be represented by the vector $B^k = \left(b_1^k, \dots, b_L^k\right),\ k = 1{,}2$. The overall quality of U_1 (IQ^1) is higher than the overall quality of U_2 (IQ^2) if $g_i^1 \geq g_i^2$ and $b_j^1 \leq b_j^2$ for all *i* and *j*. When $g_i^1 \geq g_i^2$ for some *i* but $g_i^1 < g_i^2$ for others, no conclusion can be made without further information.

Many papers adopt the method of IQA. Heshmati (2002) uses different dimensions of quality indicators to refer the teaching quality in public schools:

numbers of students with incomplete grades, average grades, spaces per students and teacher density. Fu et.al. (2017) investigates the air quality in China from different pollution indicators：$PM_{2.5}$, SO_2, and NO_X. They find a positive relationship between manufacturing sector productivity and air quality. For city level data, when the productivity increase, the pollution will reduce and improve air quality. Jun and Palacios (2016) identify 17 quality indicators are essential to determine the quality of services in mobile-banking. These 17 quality indicators are in two categories: mobile-banking application and mobile-banking customers services. Among the 17 dimensions quality indicators, they find that the mobile convenience, accuracy, various mobile application service features, ease of use, and continuous improvement are the critical quality indicators to affect the customers' satisfactory. All these papers using the IQA approach that investigates the quality indicators separately.

Another approach aggregates the dimensions stated in IQA into one single quality index to reflect the quality performance by attaching a weight to each dimension. In this approach, the examiner investigates all favourable dimensions and unfavourable dimensions of quality indicators and uses some weights to aggregate all the dimensions into one quality index to reflect the overall quality performance. I call this "*weighted quality assessment*" (WQA). Using notations in the previous example, the examiner computes a weight $\alpha_i \geq 0$ for the i_{th} favourable dimension and a weight $\beta_j \geq 0$ for the j_{th} unfavourable dimension. Then the overall quality of U_j is a function of all dimensions $WQ^j(G^j, B^j)$ as follows:

$$WQ^j\left(G^j, B^j\right) = \sum_{i=1}^{M} \alpha_i g_i^j \,/\, \sum_{i=1}^{L} \beta_i b_i^j \tag{2}$$

The overall quality of U_1, $WQ^1(G^1, B^1)$, is higher than the overall quality of U_2, $WQ^2(G^2, B^2)$, if and only if $WQ^1(G^1, B^1) > WQ^2(G^2, B^2)$.

To get the weights of different dimensions from data, one representative method is the principal components analysis (PCA). PCA uses some components to proxy the information for different variables. Each component is a linear combination of all variables. The only difference among components is that they have different weights for each variable. Each component is linearly independent of other components. The number of components is always equal to the number of variables used to construct components. The examiner only needs to pick one or several components which capture the most of data variance. Here I show the procedure of PCA to illustrate how to get the weight of WQA. The steps of the PCA are as follow based on Tipping and Bishop (1999), Suppose there are M favourable dimensions can be represented by the vector $G = (g_1, \dots, g_M)$, and the linear combination vector of G as $\mathbf{Y}$, $\mathbf{Y} = (Y_1, Y_2, \dots, Y_M)$, where $Y_i = e_{i1}G_1 + e_{i2}G_2 + \dots + e_{iM}G_M$, for $i = 1, 2, \dots, M$. Let the population variance of Y_i as $\mathrm{var}\,(Y_i) = \sum_{k=1}^{M}\sum_{l=1}^{M} e_{ik}e_{il}\sigma_{kl}$, and the population covariance of Y_i and Y_j as $\mathrm{cov}\left(Y_i, Y_j\right) = \sum_{k=1}^{M}\sum_{l=1}^{M} e_{ik}e_{jl}\sigma_{kl}$. Select the $e_{ik}\ and\ e_{il}$ to maximize the variance of Y_i,

$$\max_{e_{ik}, e_{il}} var\,(Y_i) = \sum_{k=1}^{M}\sum_{l=1}^{M} e_{ik}e_{il}\sigma_{kl} \tag{3}$$

$$\text{s.t.} \quad \sum_{j=1}^{p} e_{ij}^2 = 1$$

$$\mathrm{cov}\left(Y_i, Y_j\right) = 0 \ \text{ for } \ i \neq j$$

The linear combination vector **Y** which has maximal variance is called as *principal component*. The linear combination is sorted by percentage of explained variance λ_i, $\lambda_i = \frac{var(Y_i)}{\sum_{j=1}^{M}\sigma_{ij}^2}$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ and $\sum_{i=1}^{p}\lambda_i = 1$. If the first k principal components past the chi-square test, first k-1 principal components cannot past the test. We will choose first k principal components **Y** to replace the p dimensions of G. Similarly, the N unfavourable dimensions can be represented by the vector $B = (b_1, \dots, b_L)$, we have the linear combination vector of B as **Q**, $\mathbf{Q} = (Q_1, Q_2, \dots, Q_L)$, where $Q_i = e_{i1}B_1 + e_{i2}B_2 + \dots + e_{iL}B_L$, for $i = 1, 2, \dots, L$. The principal components are linear combinations of the original variables weighted by their contribution to explaining the variance in a particular orthogonal dimension. By sorting from the explained variance, first principal components Y_1 and Q_1 show the most important information of the favorable dimensions and unfavorable dimensions in G and B. Thus, $WQ^j(G^j, B^j) = Y_1^j / Q_1^j$

Adler and Berechman (2001) uses responses to subjective questionnaires to evaluate the performance of airports and finds the most high-quality ranking airports in West-Europe are Geneva, Milan and Munich. Their questionnaires contain five different groups of questions: suitability, reliability, using cost, satisfaction, and demand. The principal component analysis is used to aggregate the empirical data of various dimensions into a single number. Karlen et.al. (2003) review the quality indicators of soil quality assessment. There are more than twenty different quality indicators in different studies. They use PCA methods to aggregate the quality indicator or construct the soil quality function (Doran et al., 1994; Doran and Jones, 1996; Gregorich, 1996; Karlen et al., 1997). For these papers using the WQA approach, they can evaluate the relative overall quality performance for each DMU. WQA approach considers different aspects of quality indicators and can rank the relative performance of different DMUs

2.2.2 *Case 2: Production Quality Assessment*

When the favourable and unfavourable dimensions are the outputs jointly produced from utilizing resources, assessing quality must include the information of inputs and production technology. The favourable and unfavourable dimensions are called *desirable* and *undesirable outputs*, respectively. I call this "*production quality assessment*" (PQA).

One approach of PQA is to compute quantity indices of desirable and undesirable outputs. This subsection briefly describes the procedures of this approach. Let the inputs, desirable outputs and undesirable outputs be $x \in \mathbb{R}_+^N$, $g \in \mathbb{R}_+^M$ and $b \in \mathbb{R}_+^L$, respectively. The production set is $\mathfrak{I} = \{(x, g, b)\colon \text{x can produce } (g, b)\ \}$. Consider two production units, 1 and 2, which have the same production technology. Production unit *k* uses input vector x^k to production (g^k, b^k), $k = 1{,}2$. Thus, $(x^k; g^k, b^k) \in \mathfrak{I}$ for all *k*. The quality of unit *k* is reflected by desirable outputs g^k relative to undesirable outputs b^k. To compare the quality between units 1 and 2, one way is to construct a quantity index of desirable outputs $Q_g(g^1, g^2; x^r, b^r)$ of unit 1 relative to unit 2. x^r and b^r are reference vectors. Since inputs and undesirable outputs are fixed, the value of Q_g indicates the change of g^2 from g^1. When $Q_g(g^1, g^2; x^r, b^r) > 1$, there is an increase in the quantities of unit 1's desirable outputs relative to unit 2's. Similarly, a quantity index of undesirable outputs $Q_b(b^1, b^2; x^r, g^r)$ is constructed to indicate the change from b^2 to b^1. An overall quality performance index is defined as the ratio between these two quantity indices:

$$PQ_k \coloneqq \frac{Q_g(g^1,g^2;x^r,b^r)}{Q_b(b^1,b^2;x^r,g^r)}. \tag{4}$$

Unit 1 has higher quality if $PQ_1 > PQ_2$. This is the method introduced by Färe et.al (2004) to measure the environmental performance in the electricity sector.

2.3 Quality Assessment in Hospital Sector

From the discussion of Section 2.2, quality assessment has three popular methods: IQA, WQA and PQA. For treatment quality evaluation using IQA in the hospital sector, many studies use patient satisfaction as a measure of treatment quality. Aiken et.al. (2012) use patients' satisfaction as a quality proxy to investigate the determinants of satisfaction in 12 countries. They find the key factors to affect the quality is the environment of nursing services. Al-Abri and Al-Balushi (2014) make a systematic review of these papers using the patient satisfaction to assess the treatment quality (Bjertnaes, et.al., 2012; Ahmad, et.al., 2011; Clever, et.al., 2008; Cheng, et.al., 2003). They find out controversial results from different studies. The patient satisfaction is more related to the patients' age and health status rather than the treatment quality of hospitals. Fenton et al. (2012)

has investigated the relationship between treatment quality from the risk-adjusted mortality and patients' satisfaction, higher satisfaction due to the drug abuse, and cause higher mortality rate.

The second method is WQA which aggregates different objective measures (mortality rate, reinfection rate, the average number of days of hospitalisation, medical malpractice rate, etc.) into a single index to reflect treatment quality (Jha et al., 2005; Schreyögg et al., 2011; Nolan, 2001). The single dimension outcome of medical care, in terms of perinatal mortality rate and surgical fatality rates has been frequently used as an indicator of the quality of medical care. Pitches et.al. (2007) use the risk-adjusted mortality rate to evaluate the performance of treatment quality. This measure uses the mortality rate of different types of diseases in hospitals adjusted to the overall mortality rate of each disease. This measure is more objective than patient satisfaction. Donabedian (2005) reviews several measures of quality and summarizes that the measures of quality are multidimensions and objective to reflect the value of the medical care system.

The hospitals or other healthcare providers can be treated as production units. They use labour (doctors, nurses and other medical staffs) and capital (medical equipment, beds and other capital-related factors) as inputs to provide desirable outputs (e.g. inpatient treatments, outpatient treatments and other medical services) and joint-products which are undesirable (e.g., mortality, medical errors and other undesirable outcomes). Such production approach has been adopted to evaluate hospital performance based on the production process (Hollingsworth, 2003; Kiadaliri et al., 2013; Pelone et al., 2015; Rosko and Mutter, 2008). As stated in Section 2.1, the focus of this thesis is on the quality of producing desirable outputs relative to undesirable outputs. The production quality assessment is appropriate. One objective is to find an index that reflects the treatment quality of desirable outputs relative to undesirable outputs. This is discussed in the next subsection.

2.4 **The Output-oriented Environmental Performance Index**

As concluded in the previous subsection, the approach of production quality assessment is appropriate for evaluating multidimensional treatment quality when hospitals can be treated as production units. However, although production quality assessment has been applied to the hospital sector to measure productive quality, there seems no empirical works using the PQA to study treatment quality from my knowledge. It is observed that the production of successful treatments along with deaths is similar to the production of electricity along with pollutants. This section argues the environmental performance index in the field of environmental protection is also a valid measure of treatment quality.

Färe et al. (1989) describe good outputs and bad outputs as desirable outputs and undesirable outputs. They first propose assuming weak disposability when undesirable outputs appear in the production process and are jointly produced with desirable outputs. Färe et al. (2004) define environmental performance as the degree to which a production unit can produce good outputs while simultaneously accounting for the reduction of bad outputs. They construct an *environmental performance index* (EPI) to evaluate environmental performance as described below.

Let the inputs, desirable outputs and undesirable outputs be $x \in \mathbb{R}_{+}^{N},\ g \in \mathbb{R}_{+}^{M}$ and $b \in \mathbb{R}_{+}^{L}$, respectively. The production set at time $t = 1, \ldots, T$ is $\mathfrak{J}^{t} = \{(x, g, b)\colon \text{x can produce } (g, b)\ in\ time\ t\}$. Define the output set $P^{t}(x) \coloneqq \{(g, b)\colon (x, g, b) \in \mathfrak{J}^{t}\}$, the desirable output set $P^{t}(x|b) \coloneqq \{g\colon (x, g, b) \in \mathfrak{J}^{t}\}$, the undesirable output set $P^{t}(x|g) \coloneqq \{b\colon (x, g, b) \in \mathfrak{J}^{t}\}$ and the input requirement set $L^{t}(g, b) \coloneqq \{x\colon (x, g, b) \in \mathfrak{J}^{t}\}$. Desirable and undesirable outputs are jointly weakly disposable: if $(x, g, b) \in \mathfrak{J}^{t}$ and $0 \leqq \theta \leqq 1$, then $(x, \theta g, \theta b) \in \mathfrak{J}^{t}$. Desirable outputs and inputs are strongly disposable: if $(x, g, b) \in \mathfrak{J}^{t}$ and $(-x, g) \leqq (-x', g')$, then $(x', g', b) \in \mathfrak{J}^{t}$. It is obvious that,

$$(x; g, b) \in \mathfrak{I}^t \Leftrightarrow g \in P^t(x|b) \Leftrightarrow\ b \in P^t(x|g) \Leftrightarrow x \in L^t(g, b).$$

The following technological properties are assumed: $\mathfrak{I}^t$ shows constant returns to scale, i.e., for $\tau > 0$, $\tau\mathfrak{I}^t = \mathfrak{I}^t$. All inputs are essential: $x_n = 0 \Longrightarrow (g, b) = 0$. Inputs and desirable outputs are strongly disposable, i.e., $(x; g, b) \in \mathfrak{I}^t \Longrightarrow (x'; g', b) \in \mathfrak{I}^t$ for $(-x', g') \leqq (x, g)$. Desirable and undesirable outputs are jointly weakly disposable, i.e., $(x; g, b) \in \mathfrak{I}^t \Longrightarrow (x; \lambda g, \lambda b) \in \mathfrak{I}^t$ for $0 < \lambda \leqq 1$.

The distance functions for desirable outputs and undesirable outputs with respect to technology $\mathfrak{I}^t$ are

$$D_g(x, g, b|\mathfrak{I}^t) \coloneqq \underset{\theta}{inf} \left\{\theta \colon \left(x; \frac{g}{\theta}, b\right) \in \mathfrak{I}^t\right\}. \tag{5}$$

$$D_b(x, g, b|\mathfrak{I}^t) \coloneqq \sup_{\theta} \left\{\theta \colon \left(x; g, \frac{b}{\theta}\right) \in \mathfrak{I}^t\right\}. \tag{6}$$

Let (x^r, b^r) be two reference vectors for inputs and undesirable outputs. The *desirable output quantity index,* $Q_g^{out}(g^{\mathrm{t}}, g^0; b^r, x^r|\mathfrak{I}^t)$ is defined as,

$$Q_g^{out}(g^t, g^0; x^r, b^r|\mathfrak{I}^t) \coloneqq \frac{D_g(x^r; g^t, b^r|\mathfrak{I}^t)}{D_g(x^r; g^0, b^r|\mathfrak{I}^t)}. \tag{7}$$

On the right-hand side of (7), the only change is from g^0 to g^t. Hence tis index captures the change in the desirable output vector g^{t} relative to the desirable output vector g^0 with respect to technology $\mathfrak{I}^t$ and reference vectors (x^r, b^r).

Similarly, the *undesirable output quantity index* measures the change in the undesirable output vector from b^0 to b^t with respect to technology $\mathfrak{I}^t$ and reference vectors (x^r, g^r):

$$Q_b^{out}(b^t,b^0;x^r,g^r|\mathfrak{I}^t) := \frac{D_b(x^r;g^r,b^t|\mathfrak{I}^t)}{D_b(x^r;g^r,b^0|\mathfrak{I}^t)}. \qquad (8)$$

Finally, from Equations (5) and (6), the *output-oriented environmental performance index* EPI^{out} of the investigated output vectors $(g^{\mathrm{t}},b^{\mathrm{t}})$ relative to the output vectors (g^0,b^0) is

$$EPI^{out}(g^{\mathrm{t}},b^{\mathrm{t}};g^0,b^0;x^r,g^r,b^r|\mathfrak{I}^t) = \frac{Q_g^{out}(g^t,g^0;x^r,b^r|\mathfrak{I}^t)}{Q_b^{out}(b^{\mathrm{t}},b^0;x^r,g^r|\mathfrak{I}^t)}. \qquad (9)$$

When desirable outputs and undesirable outputs are produced by utilising inputs, EPI^{out} measures the units of desirable outputs quantity index per unit of undesirable outputs quantity index. Since both types of outputs are jointly produced, environmental performance in this thesis refers to more desirable outputs and less undesirable outputs. By reflecting the relative change in from the initial output vectors (g^0,b^0) to the investigated output vectors $(g^{\mathrm{t}},b^{\mathrm{t}})$, EPI^{out} can be used to measure environmental performance. When the value of EPI^{out} increases, there are more desirable outputs produced per unit of undesirable outputs. Thus, environmental performance improves when EPI^{out} increases. The EPI introduced by Färe et al. (2004) has been widely used to evaluate environmental performance in empirical studies in the energy sector (Tsoulfas and Pappis, 2008; Munksgaard et al., 2007).

The main feature of EPI is the construction of quantity indices of the two types of outputs. As long as there is a well-defined production technology, EPI is applicable. Medical services are like desirable outputs and mortalities are like undesirable outputs. With doctors, beds, etc. as inputs, the method of Färe et al. (2004) can be adopted to construct quantity indices for the good outputs and bad outputs in the hospital sector to measure treatment quality.

2.5 Modelling Production Technology with Undesirable Outputs in DEA

Undesirable outputs exist in many industries (e.g. air pollution caused by electricity power generation, deaths during medical treatments). On the one hand, this type of outputs has negative values and are not wanted. They are classified as undesirable outputs. On the other hand, these outputs are jointly produced with desirable outputs. At the end of the last subsection, it is mentioned that EPI can be adopted to measure treatment quality when the production technology of producing desirable outputs and undesirable outputs from utilizing inputs is well-defined. Thus, the first major issue of adopting the production quality assessment is the modelling of the production technology from empirical data. Yet undesirable outputs are unavoidable if positive quantities of desirable outputs are produced. The existence of joint production complicates theoretical and empirical modelling of the production technology.

The first concern is finding an empirical production frontier. There are two major approaches of estimating the production frontier from data: the data envelopment analysis (DEA) approach and the stochastic frontier analysis (SFA) approach. The former applies linear programming techniques and the frontier is deterministic. The latter incorporates random errors so that the frontier is deterministic. Econometric techniques are used. Hollingsworth (2008) systematically reviews 317 journal papers that apply the SFA and DEA approaches to productivity and efficiency measurement in the hospital sector. It finds out that using the DEA approach to investigate the productivity and efficiency in the hospital sector becomes more popular. Ferrari (2006) uses the DEA approach to measure Malmquist productivity index of 53 Scottish hospitals during the NHS reform period from 1992 to 1997. It concludes that the total factor productivity has slightly improved during the reform and this improvement comes from technological improvement. Street (2003) compares SFA and OLS in UK hospitals. It finds that the efficiency scores are sensitive to the error distribution. Kiadaliri et.al. (2013) review 29 studies using SFA and DEA for measuring the efficiency in hospitals. They find no significant different results from both approaches. In conclusion, DEA can deal with multiple outputs whereas SFA in general can only apply to a single output case. Thus, in the presence of multidimensional outputs, the DEA approach is more appropriate for quality assessment in hospital sector.

The second concern is the incorporation of undesirable outputs in the production process. Following Dakpo et al.'s (2016) classification of methods of incorporating undesirable outputs in the production process, the pros and cons of various ways to model the production technology with undesirable outputs are discussed in the following subsections.

2.5.1 Treating undesirable outputs as inputs

Some researchers observed the similarities between undesirable outputs and inputs. Both must appear in the production process. Producing more units of outputs is usually associated with larger quantities of undesirable outputs and inputs. Further, both undesirable outputs and inputs are bounded below to produce a certain output vector. In view of these, Hailu and Veeman (2001) suggested treating undesirable outputs as strongly disposable inputs in constructing the empirical production frontier.

Let the inputs, desirable outputs and undesirable outputs be $x \in \mathbb{R}_{+}^{N},\ g \in \mathbb{R}_{+}^{M}$ and $b \in \mathbb{R}_{+}^{L}$, respectively. The production set is $\mathfrak{I} = \{(x, g, b)\text{: x can produce } (g, b)\ \}$. Outputs and inputs are strongly disposable: if $(x, g, b) \in \mathfrak{I}$ and $(-x, g, -b) \leqq (-x', g', -b')$, then $(x', g', b') \in \mathfrak{I}$. There are K numbers of decision-making units (DMU). Production unit *k* uses input vector x^k to production (g^k, b^k), $k = 1,2, .., \text{K}$. Thus, $(x^k; g^k, b^k) \in \mathfrak{I}$ for all *k*, where $x^k = \left(x_1^k, \dots, x_N^k\right), g^k = \left(g_1^k, \dots, g_M^k\right)\ and\ b^k = \left(b_1^k, \dots, b_L^k\right)$. The empirical production set suggested by Hailu and Veeman (2001) is:

$$\mathfrak{I}^{HV} := \left\{\begin{matrix} (x; g, b)\text{: } \sum_{k=1}^{K} z_k\, g_m^k \geqq g_m, m = 1, \dots, M; \\ \sum_{k=1}^{K} z_k\, b_l^k \leqq b_m, l = 1, \dots, L; \\ \sum_{k=1}^{K} z_k\, x_n^k \leqq x_n, n = 1, \dots, N; \\ z_k \geqq 0, where\ k = 1, \dots, K. \end{matrix}\right\} \tag{10}$$

The third inequality is a standard way to model inputs in DEA. In this formulation, the mathematical properties of *b* and *x* are the same.

Many studies adopt the SDA to model the undesirable outputs as inputs. For example, Paul, et.al. (2002) study the costs and chemical uses in U.S. agricultural production. By treating undesirable outputs as inputs, they hypothesize the producer maximizing the desirable outputs and minimizing the inputs and undesirable outputs simultaneously. Yang and Pollitt (2009) use four-stage models to compute the environmental inefficiency of the energy sector in China. In the basic model of DEA, the undesirable output SO2 emission are modelled with SDA as the same as inputs. Mahlberg and Sahoo (2011) decompose the Luenberger productivity indicator with SDA for both undesirable outputs and inputs. It argues that undesirable outputs can be treated as inputs since both two incur cost for the producer since reducing undesirable outputs requires some additional efforts and resources to meet the environmental regulations. Firms try to minimize both inputs and undesirable outputs at the same time. Considine and Larson (2006) consider the market transfer of the pollution permit quota which is costly for firms. Undesirable outputs can be treated as the inputs since the pollution permit quota is one of the resources for producing desirable outputs.

An advantage of this approach is its simplicity. However, treating undesirable outputs as inputs cause problems. Färe and Grosskopf (2003) make critical comment to the SDA approach from Hailu and Veeman (2001). This comment shows that "the monotonicity condition introduced by Hailu and Veeman is inconsistent with physical laws." Base on the monotonicity assumption in equation (10), if the set is defined as $(x, g, b) \in \Im$ and $b' > b$, then

$(x, g, b') \in \Im$ must satisfied. Färe and Grosskopf (2003) point out "This implies that using fixed amounts of energy, capital, labor, and materials can yield an unbounded amount of undesirable outputs such as total suspended solids, etc. This is, of course, physically impossible."

Also, substitution effect is not expected to appear between undesirable outputs and inputs. If such substitution does exist, we can reduce the quantities of undesirable outputs by employing higher quantities of inputs. Since the two types of outputs are jointly produced, the substitution between undesirable outputs and inputs are in general not possible without affecting the quantities of desirable outputs. Note that in $\Im^{HV}$, such substitutions are allowed without affecting the production of desirable outputs. If the set is defined as $(x, g, b) \in \Im$, there are always exist $b' > b$ and $x' < x$ that let $(x', g, b') \in \Im$ satisfied. That implies we can always use less labour and materials to generate the same desirable outputs with more pollutants. This is obviously violating the physical rule also. There is an inconsistency between theoretical concepts and empirical modelling.

2.5.2 Material balance principal approach

The material balance principle (MBP) is introduced by Field (1994). The idea comes from the fundamental physics principle that the materials do not disappear during the transformation of outputs to outputs.

Let the inputs, desirable outputs and undesirable outputs be $x \in \mathbb{R}_+^N$, $g \in \mathbb{R}_+^M$ and $b \in R_+$, respectively. In the production approach, resources appear on the input side. Desirable outputs and residual undesirable outputs are on the output side. Under the material balance principle, the materials on the two side should be equivalent. The equation implied by the MBP as described in Coelli et.al. (2005) is: $b = N'x - M'g$, where N$\in \mathbb{R}_+^N$ and M$\in \mathbb{R}_+^M$ are non-negative shadow price vectors. There are K numbers of decision-making units (DMU). Production unit *k* uses input vector x^k to production (g^k, b^k), $k = 1,2,..,$K. Thus, $(x^k; g^k, b^k) \in \mathfrak{I}^{MB}$ for all *k*, where $x^k = (x_1^k, \dots, x_N^k), g^k = (g_1^k, \dots, g_M^k)\ and\ b^k$. The empirical frontier for the MBP is,

$$\mathfrak{I}^{MB} := \left\{ \begin{array}{c} (x; g, b): \sum_{k=1}^{K} z_k\, g_m^k \geqq g_m, m = 1, \dots, M; \\ \sum_{k=1}^{K} z_k\, x_n^k \leqq x_n, n = 1, \dots, N; \\ b = N'x - M'g \\ z_k \geqq 0, where\ k = 1, \dots, K. \end{array} \right\} \qquad (11)$$

Lauwers and Van Huylenbroeck (2003) define the environmental technical efficiency and environmental allocative efficiency based on the MBP in DEA approach. Coelli, et.al. (2005) explicitly introduces the method to estimate the iso-environmental line and iso-nutrition line from MBP. It decomposes the environmental efficiency into technical efficiency and environmental allocative efficiency.

However, some problems occur for the multidimensional undesirable outputs. The MBP's formula relies on the existence of one single undesirable output. For multidimensional undesirable outputs, the MBP formula cannot be applied. Also, the MBP assumes that the only source of reducing the undesirable outputs is from the allocation of input-mix. If there is only one input, based on the MBP, there is no way for further reduction of the undesirable output.

2.5.3 By-product modelling approach

The by-production approach (BPA) is a relative novel approach which was introduced by Murty et al. (2012). In the BPA, two independent production sets are assumed: one is the normal production set $\mathfrak{I}^{BP1}$ to generate the desirable outputs and the other one is the residual production set $\mathfrak{I}^{BP2}$ to generate the undesirable outputs. The materials inputs can separate into non-pollutant inputs and pollutant generate inputs: all inputs are used to produce desirable outputs; the only pollutant generate inputs are used to generate the residual undesirable outputs.

Let the inputs, desirable outputs and undesirable outputs be $x \in \mathbb{R}_{+}^{N}$, $g \in \mathbb{R}_{+}^{M}$ and $b \in \mathbb{R}_{+}^{L}$, respectively. x has two sub-vectors: $x_g \in \mathbb{R}_{+}^{S}$ is the input which generate desirable outputs only and $x_b \in \mathbb{R}_{+}^{N-S}$ is pollution-generating input which generate undesirable outputs b. The general technology $\Im^{BP}$ is defined as $\left(x_g, x_b; g, b\right) \in \Im^{BP}$ which is the interaction set for the first-stage production set $\left(x_g, x_b; g\right) \in \Im^{BP1}$ and second-stage production set$(\,x_b; b) \in \Im^{BP2}$. Thus, $\Im^{BP} = \Im^{BP1} \cap \Im^{BP2}$. The technology $\Im^{BP1}$ has strongly disposable inputs and desirable output as: if $\left(x_g, x_b; g\right) \in \Im^{BP1}$ and $\left(x_g, x_b; -g\right) \leqq \left(x'_g, x'_b; -g'\right)$, then $\left(x'_g, x'_b, g'\right) \in \Im^{BP1}$. The technology $\Im^{BP2}$ has strongly disposable inputs and undesirable output as, if $(\,x_b, b) \in \Im^{BP2}$ and $(-\,x'_b; b') \geqq (x_b; b)$, then $(\,x_b; b') \in \Im^{BP2}$. There are K numbers of decision-making units (DMU). DMU *k* uses input vector x^k to production (g^k, b^k), $k = 1,2,..,$K. Thus, $(x^k; g^k, b^k) \in \Im$ for all *k*, where $x^k = \left(x_g^k, x_b^k\right), x_g^k = \left(x_1^k, \dots, x_S^k\right), x_b^k = \left(x_{S+1}^k, \dots, x_N^k\right)$, $g^k = \left(g_1^k, \dots, g_M^k\right), and\ b^k = \left(b_1^k, \dots, b_L^k\right)$.

The empirical general production set in Murty et al. (2012) is,

$$\mathfrak{J}^{BP} := \left\{ \begin{array}{ll} (x_g, x_b; g, b): \sum_{k=1}^{K} z_k \, g_m^k \geqq g_m, & m = 1, \dots, M; \\ \sum_{k=1}^{K} z_k \, x_n^k \leqq x_n, & n = 1, \dots, N; \\ \sum_{k=1}^{K} u_k \, b_l^k \leqq b_l, & l = 1, \dots, L; \\ \sum_{k=1}^{K} u_k \, x_s^k \geqq x_s, & s = S+1, \dots, N; \\ z_k \geqq 0, u_k \geqq 0, & k = 1, \dots, K. \end{array} \right\} \quad (12)$$

Some recent empirical studies adopt BPA to estimate environmental efficiency. The BPA is suitable for modelling the production approach with the coexistence of non-pollution generating inputs and pollution generating inputs. For example, Arjomandi et al. (2018) adopt the BPA to find out the consistent improvement of the environmental performance of airlines in the European Union. Cui and Li (2017) use BAP to evaluate the performance of 29 global airlines. They find that these airlines cannot meet the requirement of “Carbon Neutral Growth from 2020” strategy.

When inputs cannot be separated into two distinct groups under joint production, BAP is not applicable. For example, in hospital sector, successful treatments and various types of deaths are desirable and undesirable outputs respectively. It is obviously unreasonable to assume some nurses, for example, are contributed to treat patients and other nurses are devoted to produce deaths.

2.5.4 Weak disposability assumption

The weak disposability assumption (WDA) is first introduced in Shephard (1970) to model the congestion of inputs. When there is over utilization of input factors, the production of outputs may reduce under a certain level. Färe et al. (1986, 1989) first propose using the weak disposability approach to model undesirable outputs in empirical studies. The crucial characteristics of this approach is that desirable outputs and inputs are strongly disposable whereas desirable outputs and undesirable outputs are jointly weakly disposable.

Desirable outputs and inputs are strongly disposable: if $(x, g, b) \in \Im^{WD}$ and $(-x, g) \leqq (-x', g')$, then $(x', g', b) \in \Im^{WD}$. Desirable and undesirable outputs are jointly weakly disposable: if $(x, g, b) \in \Im^{WD}$ and $0 \leqq \lambda \leqq 1$, then $(x, \lambda g, \lambda b) \in \Im$. There are K numbers of decision-making units (DMU). DMU k uses input vector x^k to produce (g^k, b^k), $k = 1,2,..,$K. Thus, $(x^k; g^k, b^k) \in \Im^{WD}$ for all k, where $x^k = \left(x_1^k, \dots, x_N^k\right), g^k = \left(g_1^k, \dots, g_M^k\right) \, and \; b^k = \left(b_1^k, \dots, b_L^k\right)$. The corresponding empirical production set is,

$$
\mathfrak{I}^{WD} := \left\{
\begin{array}{ll}
(x; g, b\,): \sum_{k=1}^{K} z_k\, g_m^k \geqq \theta g_m, & m = 1, \dots, M; \\
\sum_{k=1}^{K} z_k\, b_l^k = b_l, & l = 1, \dots, L; \\
\sum_{k=1}^{K} z_k\, x_n^k \leqq x_n, & n = 1, \dots, N; \\
z_k \geqq 0, & k = 1, \dots, K.
\end{array}
\right\} \tag{13}
$$

This approach has been applied to empirical studies extensively. Chung et al. (1997) construct the Malmquist productivity index with undesirable outputs using WDA. Färe et al. (2004), who adopt weak disposability of undesirable outputs to construct an environmental performance index (EPI) to evaluate environmental performance. Färe et.al. (2012) propose a Luenberger total factor productivity indicator with undesirable outputs using WDA.

The joint weakly disposability of desirable and undesirable outputs is the main characteristic of this approach. Together with null-jointness of undesirable outputs, the nature of joint production between the two types of outputs can be modelled. The appearance of multiple outputs can be handled easily in empirical studies by adopting the method of data envelopment analysis, as shown in the equation (13). In summary, for this thesis, the requirements of modelling the empirical production technology are fourfold: (i) The treatment of undesirable outputs must conform with economic principles. (ii) The method can handle multi-dimensions of desirable and undesirable outputs. (iii) The method can incorporate the nature of joint production between desirable and undesirable outputs. (iv) The method should applicable to hospital sector (should not violate the physical law of medical services). From the previous discussions, each method in 2.4.1 to 2.4.3 violates at least one requirement. The production technology assuming weakly disposable undesirable outputs satisfies all criteria. Therefore, this approach is adopted in this thesis. The empirical production set $\mathfrak{I}^{WD}$ has been shown to be not convex by Kuosmanen (2005). He then introduces a convex production set with WDA in undesirable outputs. Since most constructed production sets in DEA are convex, Kuosmanen’s (2005) formulation will be adopted in the empirical study in Chapter 4 (Essay 3).

2.6 Conclusion

This chapter reviews the literature on quality assessment of health care. Section 2.1 discusses the definition of quality in healthcare from WHO which includes six dimensions. Due to its importance and rareness in production quality assessment, evaluation of treatment quality is the focus of this thesis. Section 2.2 classifies quality assessment into non-production approach and production approach. Section 2.3 reviews the studies on treatment quality evaluation in the hospital sector. It is pointed out that the production approach is appropriate when hospitals can be treated as production units. Section 2.4 points out that the environmental performance index in the field of energy study can be borrowed to measure treatment quality. Section 2.5 summarizes major methods of modelling the production technology with undesirable outputs. It is concluded that assuming weakly disposable undesirable outputs and estimating the empirical production set by DEA is the most suitable for measuring treatment quality in the hospital sector.

Although quality assessment is a crucial issue for some sectors, there is no unifying framework to evaluate quality in the literature. By reviewing quality assessment of medical services, this chapter lays out a framework of analysis for measuring the quality of medical treatment. Specifically, it is argued that production quality assessment is better than other approaches of assessing treatment quality. The environmental performance index is a suitable tool in the production quality assessment. To compute the environmental performance index, weak disposability of undesirable outputs should be assumed in the modelling of the production technology.

Chapter 3

Environmental Performance Indices: Input-oriented and Malmquist Approaches

Undesirable outputs as joint-products are unavoidable in some industries, such as air pollution in the energy sector and deaths in treatment in the healthcare sector. There are many studies on environmental issues. Färe et al. (1989) propose the weak disposability approach to model undesirable outputs. Using the strong and weak disposability of outputs and inputs, they define desirable outputs and undesirable outputs in their theoretical framework. Chung et al. (1997) adopt this framework to use directional distance functions to construct the Malmquist productivity index with undesirable outputs. This essay follows the concept of environmental performance in Färe et al. (2004), who adopt weak disposability of undesirable outputs to construct an environmental performance index (EPI) to evaluate environmental performance. In the last chapter, EPI was suggested to be adopted for measuring treatment quality.

Although EPI has been used to evaluate environmental performance in many empirical studies in the energy sector (Tsoulfas and Pappis, 2008; Munksgaard et al., 2007), it has some unsolved problems. The current formulation of EPI is a ratio between two quantity indices of desirable and undesirable outputs constructed from output distance functions. Each quantity index indicates the change from a base vector to a current vector. From the formula of (?), the computation of EPI involves benchmark production frontiers and reference vectors. Surprisingly, all existing EPI applications use minimum observed outputs and maximum observed inputs as reference vectors and base vectors. Further, all observations from different periods are evaluated with respect to the same production frontier. The reference vectors and production frontier used in the EPI are not common in other indices of the data envelopment analysis approach. For example, the Malmquist productivity index and the Hicks–Moorsteen productivity index use observations from the previous year or the current year as reference vectors and the previous year’s frontier or the current year’s frontier as reference frontiers to measure quantity and productivity (Kerstens et al., 2014; Tone, 2004; Färe et al., 2011). The choice of reference vectors and frontiers reflects the benchmark for the change in environmental performance. No other index uses minimum outputs and maximum inputs as reference vectors, except the EPI. Allowing different production frontiers in different years and more flexible reference vectors are the directions of research in this chapter.

The current formulation of EPI is a ratio between two quantity indices of desirable and undesirable outputs constructed from output distance functions. Its computation involves benchmark production frontiers and reference vectors. Current empirical studies adopting EPI unanimously choose restrictive versions of such benchmark production frontiers and reference vectors. By pointing out the limitations of using output distance functions, this chapter explores alternative measures of environmental performance in two directions. First, since output distance functions cannot ensure the existence of solutions in the computation of EPI, input distance functions are adopted to construct output quantity indices of desirable and undesirable outputs. Its ratio becomes a modified version of EPI. Second, the original formulation of EPI is a ratio between two quantity indices. The environmental performance index is similar to a productivity index. So, the Malmquist productivity index is borrowed to capture the “productivity” between desirable and undesirable outputs. One main contribution of this chapter is the construction of two new formulas of EPI from input distance functions and the Malmquist productivity index. These two new formulas improve upon the existing one in the sense of allowing conventional benchmark production frontiers and reference vectors in the computation of environmental performance.

The rest of this chapter is organised as follows. Section 3.1 provides vigorous evidence of the weaknesses of the current output-oriented EPI. Section 3.2 introduces the input-oriented EPI. Section 3.3 presents the Malmquist EPI and its decomposition. Section 3.4 discusses the chosen reference vectors and frontiers. Finally, Section 3.5 concludes the essay.

3.1 The Weaknesses of the Output-oriented EPI

The current formulation of EPI is a ratio between two quantity indices of desirable and undesirable outputs constructed from output distance functions. Its computation involves benchmark production frontiers and reference vectors. As stated in the introduction of this chapter, current empirical studies adopting EPI unanimously choose restrictive versions of such benchmark production frontiers and reference vectors. The key reason is that the output-oriented EPI has restrictive conditions to ensure feasibility. Previous studies applying the EPI do not mention its feasible conditions. This subsection studies this issue in detail.

Since distance functions are used to compute EPI, the condition of having finite values in these functions is crucial. This is explored in the following proposition.

Proposition 3.1 Let $\Im$ be a production set with constant returns to scale. Suppose that inputs and desirable outputs are strongly disposable, while undesirable outputs are weakly disposable. Suppose that $D_g(x, g, b|\Im^t)$ is a finite number for $(x, g, b) \in \Im^t$. Consider the reference vectors $x^r \in \mathbb{R}^N_{++}$, $g^r \in \mathbb{R}^M_{++}$ and $b^r \in \mathbb{R}^L_{++}$. Let $(x^t, g^t, b^t) \in \Im^t$.

a. If $(x^r, g, b^r) \in \Im^t$ for some $g > 0$, then $D_g(x^r, g^t, b^r|\Im^t)$ is a finite number.

b. Define a cone: $K(x^r, g^r) \coloneqq \{b > 0 : \alpha b \in P^t(x^r|g^r)\ for\ some\ \alpha > 0\}$. Then $D_b(x^r, g^r, b^t|\Im^t)$ is a finite number if and only if $b^t \in K(x^t, g^t)$.

Proposition 3.1a suggests that the output distance function for desirable outputs is a finite number if the undesirable output vector b^r can be produced by x^r for a vector with positive desirable outputs. Thus, this distance function is Well defined. Conversely, although (x^t, g^t, b^t) are feasible, the undesirable output vector b^t needs not be in the cone $K(x^r, g^r)$ under weak disposability assumption. The following results are immediate consequences of Proposition 3.1.

Corollary to Proposition 3.1a Suppose that desirable outputs are strongly disposable. Let x^r and b^r be two reference vectors for inputs and undesirable outputs. If there is $g > 0$ such that $(x^r; g,\ b^r) \in \Im^t$ and $(x^r; g,\ b^r) \in \Im^{t-1}$, $Q_g^{out}(g^t, g^{t-1}\ ;\ x^r,\ b^r)$ is a finite number for any g^{t-1} and g^t.

Proof:

From Equation (7),

$$Q_g^{out}(g^t, g^{t-1}; x^r, b^r) := \frac{D_g(x^r; g^t, b^r|\Im^t)}{D_g(x^r; g^{t-1}, b^r|\Im^{t-1})}. \tag{14}$$

As $(x^r; g,\ b^r) \in \Im^*$ for some $g > 0$, I have $\frac{g^t}{\theta} < g$ for any g^t, *t = 1, ..., T*, if $\theta > 0$ is sufficiently large. As desirable outputs are strongly disposable, I have $\left(x^r; \frac{g^t}{\theta},\ b^r\right) \in \Im^t$. Applying Proposition 3.1a, let $g^* = \frac{g^t}{\theta^*} \geqq g$, so that $(x^r; g^*,\ b^r) \in \Im^t$ and $(x^r; g^* + \varepsilon,\ b^r) \notin \Im^t$ for any $\varepsilon > 0$. Then $D_g(x^r, g^t, b^r|\Im^t)$ is a finite number. Similarly, $D_g(x^r, g^{t-1}, b^r|\Im^{t-1})$ is a finite number. The value of $Q_g^{out}(g^t, g^{t-1}; x^r, b^r)$ is the ratio of two real numbers and is therefore a finite number.

Q.E.D.

Corollary to Proposition 3.1b Suppose that the desirable outputs and undesirable outputs are jointly Weakly disposable. Let x^r and g^r be two reference vectors for inputs and desirable outputs. If $(x^r;\ g^r, b^t) \in \Im^t$ for $t = 1, 2, \dots, T$, $Q_b^{out}(b^t, b^{t-1}\ ;\ x^r,\ b^r)$ is a finite number.

Proof:

From Equation (8),

$$Q_b^{out}(b^t, b^{t-1}; x^r, g^r) := \frac{D_b(x^r; g^r, b^t)}{D_b(x^r; g^r, b^{t-1})}. \tag{15}$$

As $(x^r;\ g^r, b^t) \in \Im^t$ and $(x^r;\ g^r, b^{t-1}) \in \Im^{t-1}$. From Proposition 3.1b, the undesirable output vectors b^t and b^{t-1} are in the cone $K(x^r, g^r)$, with $K(x^r, g^r) \coloneqq \{b^t > 0 \colon \alpha b^t \in P^t(x^r|g^r)\ for\ some\ \alpha > 0\}$. As $b^t \in K(x^t, g^t)$, $D_b(x^r, g^r, b^t|\Im^t)$ is a finite number. As $b^{t-1} \in K(x^{t-1}, g^{t-1})$, $D_b(x^r, g^r, b^{t-1}|\Im^t)$ is a finite number. The value of $Q_b^{out}(g^t, g^{t-1}; x^r, b^r)$ is the ratio of two real numbers and is therefore a finite number.

Q.E.D

As the output-oriented EPI is constructed from the quantity indices of the desirable and undesirable outputs, if both quantity indices are finite numbers, the EPI score is a finite number. An ideal property of this index is that the EPI scores should not change under different reference vectors. In other words, a good environmental performance index should be independent of the reference vectors. This is a demanding condition, as shown in the following proposition.

Proposition 3.2 Consider the desirable output vectors g^{t-1} and g^t and the undesirable output vectors b^{t-1} and b^t. For any reference vector $\left(x_\varphi^r; g_\varphi^r, b_\varphi^r\right)$ and $(x_\omega^r; g_\omega^r, b_\omega^r)$,

$$EPI^{out}\left(g^t, b^t; g^{t-1}, b^{t-1}; x_\varphi^r, g_\varphi^r, b_\varphi^r\right)$$

$$= EPI^{out}(g^t, b^t; g^{t-1}, b^{t-1}; x_\omega^r, g_\omega^r, b_\omega^r) \qquad (16)$$

if and only if for any $(x; g, b)$

$$D_g(x; g, b) = f(g) \cdot D_g(x; 1, b) \qquad (17)$$

and

$$D_b(x; g, b) = h(b) \cdot D_b(x; g, 1). \qquad (18)$$

Proof:

Suppose that Equations (17) and (18) hold. By combining these two equations with Equations (5) and (6),

$$EPI^{out}\left(g^t, b^t; g^{t-1}, b^{t-1}; x_\varphi^r, g_\varphi^r, b_\varphi^r\right)$$
$$= \frac{f(g^t)\cdot D_g\left(x_\varphi^r; 1, b_\varphi^r\right)}{f(g^{t-1})\cdot D_g\left(x_\varphi^r; 1, b_\varphi^r\right)} \Big/ \frac{h(b^t)\cdot D_g\left(x_\varphi^r, g_\varphi^r, 1\right)}{h(b^{t-1})\cdot D_g\left(x_\varphi^r, g_\varphi^r, 1\right)}$$
$$= \frac{f(g^t)\cdot h(b^{t-1})}{f(g^{t-1})\cdot h(b^t)}.$$

As the reference vectors $\left(x_\varphi^r; g_\varphi^r, b_\varphi^r\right)$ do not appear on the right side, the value of EPI^{out} for $(g^t, b^t; g^{t-1}, b^{t-1})$ will be the same regardless of the reference vectors. This means that

$$EPI^{out}\left(g^t, b^t; g^{t-1}, b^{t-1}; x_\varphi^r, g_\varphi^r, b_\varphi^r\right) = EPI^{out}(g^t, b^t; g^{t-1}, b^{t-1}; x_\omega^r, g_\omega^r, b_\omega^r).$$

Conversely, suppose that Equation (16) holds. From Equation (9), I have

$$\frac{Q_g\left(g^t, g^{t-1}; x_\varphi^r, b_\varphi^r\right)}{Q_b\left(b^t, b^{t-1}; x_\varphi^r, g_\varphi^r\right)} = \frac{Q_g(g^t, g^{t-1}; x_\omega^r, b_\omega^r)}{Q_b(b^t, b^{t-1}; x_\omega^r, g_\omega^r)}. \tag{19}$$

Given that (g^t, g^{t-1}) do not appear in the denominators and (b^t, b^{t-1}) do not appear in the numerators, it is obvious that

$$Q_g\left(g^t, g^{t-1}; x_\varphi^r, b_\varphi^r\right) = \gamma \cdot Q_g(g^t, g^{t-1}; x_\omega^r, b_\omega^r) \tag{20}$$

$$Q_b\left(b^t, b^{t-1}; x_\varphi^r, g_\varphi^r\right) = \gamma \cdot Q_b^{out}(b^t, b^{t-1}; x_\omega^r, g_\omega^r), \tag{21}$$

where $\gamma > 0$.

As Färe and Grosskopf (2004) p.24 point out, Equations (20) and (21) hold for any $\left(x_\varphi^r; g_\varphi^r, b_\varphi^r\right)$ and $(x_\omega^r; g_\omega^r, b_\omega^r)$ if and only if Equations (17) and (18) are true. This completes the proof.

Q.E.D

Proposition 3.2 shows the condition of independence of the reference vectors. However, these two conditions are difficult to satisfy in the distance functions for desirable outputs and undesirable outputs. Thus, the reference vectors will affect the evaluation of environmental performance. The current output-oriented EPI approach in empirical applications uses unconventional reference vectors and benchmark production frontiers. Let the observed data of desirable outputs, undesirable outputs and inputs for firm k at time t, $k = 1, \ldots, K$ and $t = 1, \ldots, T$, be $g_k^t = (g_{k1}^t, \cdots, g_{kM}^t)$, $b_k^t = (b_{k1}^t, \cdots, b_{kL}^t)$ and $x_k^t = (x_{k1}^t, \cdots, x_{kN}^t)$, respectively. Current studies adopting EPI^{out} unanimously use one grand frontier $\mathfrak{J}^*$ for the entire study period, i.e., $\mathfrak{J}^1 \cup \mathfrak{J}^2 \cup \ldots \cup \mathfrak{J}^T = \mathfrak{J}^*$ such that $(x_k^t, g_k^t, b_k^t) \in \mathfrak{J}^*$ for all $k = 1, \ldots, K$ and t. The grand frontier can be interpreted as a reference technology consisting of all observations for the years in the sample, as proposed by Pastor and Lovell (2005). When the data are updated, the grand frontier is likely to shift. The EPI score will change if an additional year of observations is added to the sample. That is, all values of the EPI in previous years may be different when future data are included, which is its main drawback. Consider the following vectors,

$$g_m^{min} \coloneqq \min_k \{g_{km}^t : k = 1, \cdots, K\}, m = 1, \cdots, M.$$

$$x_n^{max} := \max_k \{x_{kn}^t : k = 1, \cdots, K\}, n = 1, \cdots, N.$$

In existing empirical applications of EPI, the reference vectors are $(g^r, b^r, x^r) = \left(g^{min}, b^r, x^{max}\right)$, where b^r is arbitrary. I call $\left(g^{min}, b^r, x^{max}\right)$ *min-max reference vectors*. Based on the corollaries of Proposition 3.1a and 3.1b, it is clear that,

$(x^{max}, g, b^r) \in \mathfrak{I}^*$ for some $g > 0$ and $K\left(x^{max}, g^{min}\right) := \left\{b > 0 \colon \alpha b \in P^*\left(x^{max} | g^{min}\right) \, for \; some \; \alpha > 0\right\}$.

From corollaries of Proposition 3.1, it is clear that EPI^{out} is always a finite number for any observed input-output vector with respect to the min-max reference vectors and the grand frontier. I think that the following properties should be held by any index that measure environmental performance with respect to the production technology. First, the index should capture the shifts of the production frontier over time. Second, the index should allow for various choices of reference vectors.

Current EPI^{out} applications choose the min-max reference vectors with the grand frontier. Then, $EPI^{out}\left(x_k^t;\ g_k^t, b_k^t; x^{max}, g^{min},\ b^r\right)$ is well-defined for $k = 1, \cdots, K$ and $t = 1, \ldots, T$. If the min-max reference vectors and the grand frontier are not adopted simultaneously, EPI^{out} may have no solution, as shown in the corollary to Proposition 3.1. Here is an example of $b \notin K(x^r, g^r)$ in Graph 1. Suppose that there are two periods $t1$ and $t2$. In each period, there are three firms $f1$, $f2$ and $f3$. In the production process, there is only one desirable output g, one input x and two undesirable outputs $b1$ and $b2$. The quantity of inputs and outputs are listed in table 1,

Table 1: Numerical example of infeasible case

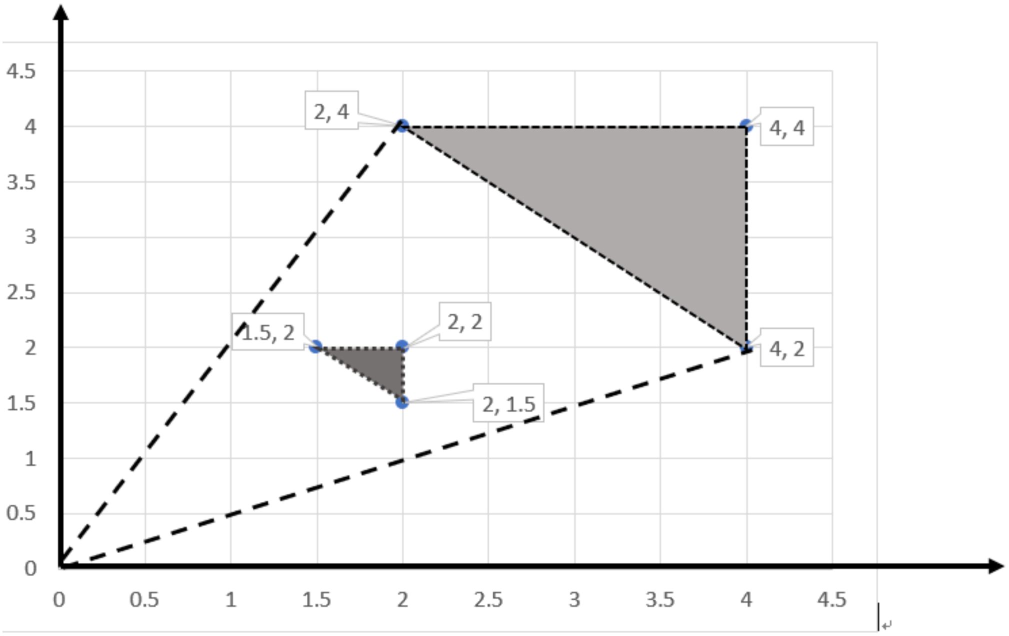
Time	PU	g	x	$b1$	$b2$
$t1$	$f1$	5	5	1.5	2
$t1$	$f2$	5	5	2	1.5
$t1$	$f3$	5	5	2	2
$t2$	$f1$	10	10	2	4
$t2$	$f2$	10	10	4	2
$t2$	$f3$	10	10	4	4

I can conclude that EPI^{out} is not feasible as

$$b \notin K(x^r, g^r) := \{b > 0 : \alpha b \in P^*(x^r = 5 | g^r = 5)\ for\ some\ \alpha > 0\}.$$

As shown in the following graph,

Graph 1



The min-max reference vectors are one of many choices of reference vectors. The grand frontier is only one of many choices of benchmark production frontier. The above example shows that the output-oriented EPI may well-defined. Therefore, the two properties stated before are not satisfied. There are additional drawbacks. The grand frontier may shift when new data are added. Thus, existing results will be affected by future data, which is not satisfactory. Furthermore, the min-max reference vectors are extreme values of desirable outputs and inputs, the values of EPI computed based on them are very sensitive to outliers. In summary, the current formulation of EPI allows for very restrictive benchmark production frontier and reference vectors only. A formula that can incorporate flexible choices is needed. The next section explores a new direction from input-orientation.

3.2 Input-oriented Environmental Performance Index

This section proposes an alternative input-oriented approach to construct the environmental performance index. To proceed, let us define the *production set with unrestricted undesirable outputs* as $\Im_b^t \coloneqq \{(x,g):(x;g,b) \in \Im^t, b \geqq 0\}$. In this production set, (x,g) are feasible as long as (g,b) can be produced by x for some $b \geqq 0$. Similarly, the *production set with unrestricted desirable outputs* is $\Im_g^t \coloneqq \{(x,b):(x;g,b) \in \Im^t, g \geqq 0\}$.

The *input distance function for desirable outputs* with respect to $\Im_b^t$ is

$$D_i(x,g|\Im_b^t) \coloneqq \sup_{\lambda}\left\{\lambda:\left(\frac{x}{\lambda},g\right) \in \Im_b^t\right\}. \tag{22}$$

This is a modified version of the input-oriented distance function. When the function is defined, there is a vector of undesirable outputs $b \in \mathbb{R}_+^L$ such that $\left(x/D_g^{in}(x;\ g|\Im_b^t);g,b\right) \in \Im^t$. Similarly, the *input distance function for undesirable outputs* with respect to $\Im_g^t$ is

$$D_i\left(x,b|\Im_g^t\right) \coloneqq \sup_{\lambda}\left\{\lambda:\left(\frac{x}{\lambda},b\right) \in \Im_g^t\right\}. \tag{23}$$

Let x^r be a reference vector of inputs. The *input-oriented quantity index of the desirable output vector* g^{t} *relative to the vector* g^{t-1} *with respect to the technology in* t is

$$Q_g^{in}(g^{\mathrm{t}},g^{t-1};x^r|\Im_b^t) = \frac{D_i(\ g^{t-1};x^r|\Im_b^t)}{D_i(\ g^t;x^r|\Im_b^t)}. \tag{24}$$

Note that under strong disposability of inputs, the input distance function for desirable outputs decreases Weakly in g. If $g^{t-1} \leqq g^{t}$, then $D_i(g^{t-1}; x^r | \Im_b^t) \geqq D_i(g^{t}; x^r | \Im_b^t)$ and $Q_g^{in}(g^{\mathrm{t}}, g^{t-1}; x^r | \Im_b^t) \geqq 1$.

Similarly, the *input-oriented quantity index of the undesirable output vector* b^{t} *relative to the vector* b^{t-1} *with respect to the technology in* t is

$$Q_b^{in}\left(b^{\mathrm{t}}, b^{t-1}; x^r | \Im_g^t\right) = \frac{D_i\left(b^0; x^r | \Im_g^t\right)}{D_i\left(b^t; x^r | \Im_g^t\right)}. \tag{25}$$

The input distance function for undesirable outputs decreases Weakly in b. If $b^0 \leqq b^t$, then $D_i\left(b^0; x^r | \Im_g^t\right) \geqq D_i\left(b^t; x^r | \Im_g^t\right)$ and $Q_b^{in}\left(b^{\mathrm{t}}, b^0; x^r | \Im_g^t\right) \geqq 1$.

Finally, the *input-oriented environmental performance index of the output vectors* $(g^{\mathrm{t}}, b^{\mathrm{t}})$ *relative to the output vectors* (g^{t-1}, b^{t-1}) *with respect to the input reference vector* x^r *and the technology in* t is

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}; g^{t-1}, b^{t-1}; x^r) = \frac{Q_g^{in}(g^{\mathrm{t}}, g^{t-1}; x^r | \Im_b^t)}{Q_b^{in}\left(b^{\mathrm{t}}, b^{t-1}; x^r | \Im_g^t\right)}. \tag{26}$$

Compared with the output-oriented EPI in Equation (9) that has three reference vectors, there is only one reference vector for the input-oriented EPI in Equation (26). The main difference is that the input-oriented EPI uses an unrestricted desirable output production set or an unrestricted undesirable output production set and the input-oriented distance function to solve the infeasibility problem. To show the quantity change of undesirable outputs, all changes in the index must come from the change of undesirable outputs. I set the inputs and desirable outputs as identical reference vectors in the distance functions.

Therefore, the only difference between the reference vectors and the evaluation comes from undesirable outputs. The undesirable output quantity index is the ratio of the input-distance function of observed undesirable outputs to the reference of undesirable outputs. Thus, the quantity index shows the change in the quantity of undesirable outputs. Similarly, I can use the input-oriented distance function to show the change in the quantity of desirable outputs. The feasible conditions of unrestricted input-distance functions are the following.

Proposition 3.3 Let $\mathfrak{I}^t$, $t = 1, \cdots, T$, be a production set with constant returns to scale at time t. Suppose that inputs and desirable outputs are strongly disposable, while undesirable outputs are Weakly disposable,

a. If there is $g' > 0$ such that $(x^g; g', b^g) \in \mathfrak{I}^t$ for some $(x^g, b^g) \in \mathbb{R}_+^{N\times M}$, then the input distance function for desirable outputs with respect to $\mathfrak{I}_b^t$, $D_i(x^r, g|\mathfrak{I}_b^t)$, is well-defined for $g > 0$ and $x^r > 0$.

b. If there is $b' > 0$ such that $(x^b; g^b, b') \in \mathfrak{I}^t$ for some (x^b, g^b), then the input distance function for undesirable outputs with respect to $\mathfrak{I}_g^t$, $D_i\left(x^r, b|\mathfrak{I}_g^t\right)$, is Well defined for $b > 0$ and $x^r > 0$.

Proof:

a. Let $g > 0$, based on the hypothesis in (a), $\left(x^g; g^{'}, b^g\right) \in \mathfrak{I}^t$ for some

$(x^g, b^g) \in \mathbb{R}_+^{N\times M}$ and $g^{'} > 0$. Under constant returns to scale, $\left(\theta x^g; \theta g^{'}, \theta b^g\right) \in \mathfrak{I}^t$ for $\theta > 0$. If θ is sufficiently large, then $\theta g^{'} \geqq g$. Under strong disposability of outputs, $(\theta x^g; g, \theta b^g) \in \mathfrak{I}^t$. For any $x^r > 0$, $\tau x^r \geqq \theta x^g$ for a sufficiently large τ. Under strong disposability of inputs,

$(\theta x^g; g, \theta b^g) \in \mathfrak{I}^t \Longrightarrow (\tau x^r, g, \theta b^g) \in \mathfrak{I}^t \Longrightarrow (\tau x^r, g) \in \mathfrak{I}_g^t$. Thus, $\lambda = 1/\tau$ is a feasible solution, so that $(x^r/\lambda, g) \in \mathfrak{I}_b^t$. It follows that $D_i(x^r, g|\mathfrak{I}_b^t) \coloneqq \sup_{\lambda}\left\{\lambda: \left(\frac{x}{\lambda}, g\right) \in \mathfrak{I}_b^t\right\}$ is a finite number.

b. Let $(x^b; g^b, b) \in \mathfrak{I}^t$. For any $x^r > 0$, $\tau x^r \geqq x^b$ for a sufficiently large τ. Under strong disposability of inputs, $(x^b; g^b, b) \in \mathfrak{I}^t \Longrightarrow (\tau x^r, g^b, b) \in \mathfrak{I}^t \Longrightarrow (\tau x^r, b) \in \mathfrak{I}_b^t$. Thus, $D_i\left(x^r, b|\mathfrak{I}_g^t\right) \coloneqq \sup_{\lambda}\left\{\lambda: \left(\frac{x^r}{\lambda}, b\right) \in \mathfrak{I}_g^t\right\}$ is a finite number.

Q.E.D

Proposition 3.5 suggests that the input distance function for desirable outputs and undesirable outputs are finite numbers if the positive desirable output vector g' can be produced by the given inputs and undesirable outputs vector (x^g, b^g) and the positive undesirable output vector b' can be produced by the given inputs and desirable outputs vector (x^b, g^b). Thus, these input distance functions of desirable outputs and undesirable outputs are Well defined. The feasible conditions of the input-oriented EPI can be satisfied by the following reference vectors and frontier,

Sequential frontier: $\mathfrak{I}^{st} := \mathfrak{I}^1 \cup \mathfrak{I}^2 \cup \ldots \cup \mathfrak{I}^t$ for $t = 1, \cdots, T$.

Prior-lag references: $(x^r;\ g^r, b^r) = (x^t;\ g^t, b^t)$ or $(x^{t-1};\ g^{t-1}, b^{t-1})$.

The sequential frontier and prior-lag reference vectors are commonly used references in previous studies of the productivity index (Shestalova, 2003; Oh and Heshmati, 2010). The sequential frontier contains all observations from the first period to the current period. Thus, the sequential frontier is likely to shift up over time. The last sequential frontier is the same as the grand frontier. The prior-lag reference vectors are the most commonly used reference vectors in empirical studies of productivity

Many studies on the dynamic analysis of changes over time adopt moving reference vectors and take the geometric mean. My input-oriented environmental performance index can accommodate this. Consider the change in environmental performance from $(x^{t-1}, g^{t-1}, b^{t-1})$ to (x^{t}, g^{t}, b^{t}). I use x^{t-1} and x^{t} as reference vectors and define the *input-oriented environmental performance index with respect to the technology in* t as

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}; g^{\mathrm{t-1}}, b^{\mathrm{t-1}}; x^{t-1}, x^{t})$$
$$= \left(EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}; g^{\mathrm{t-1}}, b^{\mathrm{t-1}}; x^{t-1}) EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}; g^{\mathrm{t-1}}, b^{\mathrm{t-1}}; x^{t})\right)^{0.5}.$$

The following proposition demonstrates the feasibility of the input-oriented EPI with the sequential frontier and prior-lag reference vectors.

Proposition 3.4 Suppose that the technology sets $\mathfrak{I}_g^t$ and $\mathfrak{I}_b^t$ show constant returns to scale. Suppose that $\mathfrak{I}_g^{t-1} \subset \mathfrak{I}_g^t$ and $\mathfrak{I}_b^{t-1} \subset \mathfrak{I}_b^t$, $t = 2, \cdots, T$. Then $EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}; g^{\mathrm{t-1}}, b^{\mathrm{t-1}}; x^{t-1}, x^{t})$ is a finite number for all t.

Proof: given $\mathfrak{I}_g^{t-1} \subset \mathfrak{I}_g^t$, I have $(x^i, b^i) \in \mathfrak{I}_g^t, i = t-1, t$. Given $\mathfrak{I}_b^{t-1} \subset \mathfrak{I}_b^t$, I have $(x^i, g^i) \in \mathfrak{I}_b^t, i = t-1, t$. Thus, the hypotheses of Propositions 3.3a and 3.3b are satisfied. All input-oriented distance functions in Equation (19) and (20) are Well defined. Therefore, $EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}; g^{\mathrm{t-1}}, b^{\mathrm{t-1}}; x^{t-1}, x^{t})$ is a finite number.

Q.E.D.

Proposition 3.4 shows that the input-oriented EPI can guarantee a feasible solution of the sequential frontier with the prior-lag reference vectors. This is an improvement over the output-oriented EPI that offers a feasible solution only when using the min-max reference vectors with the grand frontier. My new approach is better than that of the output-oriented EPI as it allows for more choices of the benchmark technology and reference vectors. The input-oriented EPI satisfies the following desirable properties as reported in Färe et al. (2004, p.346).

a. **Homogeneity:**

$$EPI^{in}(\theta g^{\mathrm{t}}, \lambda b^{\mathrm{t}}, g^0, b^0; x^r) = \frac{\theta}{\lambda} EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r)$$

$$\text{for } \theta > 0 \text{ and } \lambda > 0.$$

b. **Time-reversal:**

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r)\, EPI^{in}(g^0, b^0, g^t, b^t; x^r) = 1.$$

c. **Transitivity:**

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^s, b^s; x^r)\, EPI^{in}(g^{\mathrm{s}}, b^{\mathrm{s}}, g^0, b^0; x^r) = EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r).$$

d. **Dimensionality:**

$$EPI^{in}(\theta g^{\mathrm{t}}, \lambda b^{\mathrm{t}}, \theta g^0, \lambda b^0; x^r) = EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r)$$

$$\text{for } \theta > 0 \text{ and } \lambda > 0.$$

Proposition 3.5 (Homogeneity) Suppose that the technology sets $\mathfrak{I}_g^t$ and $\mathfrak{I}_b^t$ show constant returns to scale. For $\tau > 0$,

$$D_i(x; \tau g) = \tau^{-1} D_i(x; g) \tag{27}$$

and

$$D_i(x;\tau b) = \tau^{-1} D_i(x;b). \tag{28}$$

Proof:

For $\tau > 0$,

$$D_i(x;\tau g) \coloneqq \sup_{\lambda}\left\{\lambda\colon \left(\frac{x}{\lambda};\tau g, b\right) \in \Im^t, b \geqq 0\right\}$$

$$= \sup_{\lambda}\left\{\lambda\colon \left(\frac{x}{\tau\lambda}; g, \frac{b}{\tau}\right) \in \frac{1}{\tau}\Im^t, \frac{b}{\tau} \geqq 0\right\}$$

$$= \frac{1}{\tau}\sup_{\tau\lambda}\left\{\tau\lambda\colon \left(\frac{x}{\tau\lambda}; g, \frac{b}{\tau}\right) \in \frac{1}{\tau}\Im^t, \frac{b}{\tau} \geqq 0\right\}$$

$$= \tau^{-1} D_i(x;g).$$

Thus, I have $D_i(x;\tau g) = \tau^{-1} D_i(x;g)$.

$$D_i(x;\tau b) \coloneqq \sup_{\lambda}\left\{\lambda\colon \left(\frac{x}{\lambda}; g, \tau b\right) \in \Im^t, g \geqq 0\right\}$$

$$= \sup_{\lambda}\left\{\lambda\colon \left(\frac{x}{\tau\lambda}; \frac{g}{\tau}, b\right) \in \frac{1}{\tau}\Im^t, \frac{g}{\tau} \geqq 0\right\}$$

$$= \frac{1}{\tau}\sup_{\tau\lambda}\left\{\tau\lambda\colon \left(\frac{x}{\tau\lambda}; \frac{g}{\tau}, b\right) \in \frac{1}{\tau}\Im^t, \frac{g}{\tau} \geqq 0\right\}$$

$$= \tau^{-1} D_i(x;b).$$

Thus, I have $D_i(x;\tau b) = \tau^{-1} D_i(x;b)$.

Q.E.D.

The following properties of quantity indices are the natural consequences of the two corollaries of Proposition 3.5.

Corollary to Proposition 3.5a Suppose that the technology sets $\Im_g^t$ and $\Im_b^t$ show constant returns to scale. For $\alpha, \beta > 0$ and $t = 1, \ldots, T$,

$$Q_g^{in}(\alpha g^{\mathrm{t}}, \beta g^{t-1}; x^r | \Im_b^t) = \frac{\alpha}{\beta} Q_g^{in}(g^t, g^{t-1}; x^r | \Im_b^t) \tag{29}$$

and

$$Q_b^{in}\left(\alpha b^{\mathrm{t}}, \beta b^{t-1}; x^r | \mathfrak{I}_g^t\right) = \frac{\alpha}{\beta} Q_b^{in}\left(b^t, b^{t-1}; x^r | \mathfrak{I}_g^t\right). \qquad (30)$$

Corollary to Proposition 3.5b Suppose that the technology sets $\mathfrak{I}_g^t$ and $\mathfrak{I}_b^t$ show constant returns to scale. For $\alpha, \beta, \gamma, \delta > 0$ and $t = 1, \dots, T$,

$$EPI^{in}(\alpha g^t, \beta b^t; \gamma g^{t-1}, \delta b^{t-1}; x^r) = \frac{\alpha\delta}{\beta\gamma} EPI^{in}(g^t, b^t; g^{t-1}, b^{t-1}; x^r).$$

Proposition 3.6 (Time-reversal) Suppose that the technology sets $\mathfrak{I}_g^t$ and $\mathfrak{I}_b^t$ show constant returns to scale. For a given reference vector x^r, I have

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r)\, EPI^{in}(g^0, b^0, g^t, b^t; x^r) = 1.$$

Proof:

From Equations (24) and (25), I have，

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r) = \frac{D_i(\, g^0; x^r | \mathfrak{I}_b^t)}{D_i(g^t; x^r | \mathfrak{I}_b^t)} \Big/ \frac{D_i\left(\, b^0; x^r | \mathfrak{I}_g^t\right)}{D_i\left(b^t; x^r | \mathfrak{I}_g^t\right)}. \qquad (31)$$

$$EPI^{in}(g^0, b^0, g^t, b^t; x^r) = \frac{D_i(\, g^t; x^r | \mathfrak{I}_b^t)}{D_i(g^0; x^r | \mathfrak{I}_b^t)} \Big/ \frac{D_i\left(\, b^t; x^r | \mathfrak{I}_g^t\right)}{D_i\left(b^0; x^r | \mathfrak{I}_g^t\right)}. \qquad (32)$$

For the same x^r, obviously, I have,

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r)\, EPI^{in}(g^0, b^0, g^t, b^t; x^r) = 1.$$

Q.E.D

Proposition 3.7 (Transitivity) Suppose that the technology sets $\mathfrak{I}_g^t$ and $\mathfrak{I}_b^t$ show constant returns to scale. For a given reference vector x^r, I have

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^s, b^s; x^r)\, EPI^{in}(g^{\mathrm{s}}, b^{\mathrm{s}}, g^0, b^0; x^r) = EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^0, b^0; x^r).$$

Proof:

From Equations (24) and (25), I have，

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^{s}, b^{s}; x^{r}) = \frac{D_i(\, g^{s}; x^{r}|\Im_b^t)}{D_i(g^{t}; x^{r}|\Im_b^t)} / \frac{D_i\big(\, b^{s}; x^{r}|\Im_g^t\big)}{D_i\big(b^{t}; x^{r}|\Im_g^t\big)}. \qquad (33)$$

$$EPI^{in}(g^{\mathrm{s}}, b^{\mathrm{s}}, g^{0}, b^{0}; x^{r}) = \frac{D_i(\, g^{0}; x^{r}|\Im_b^t)}{D_i(g^{s}; x^{r}|\Im_b^t)} / \frac{D_i\big(\, b^{0}; x^{r}|\Im_g^t\big)}{D_i\big(b^{s}; x^{r}|\Im_g^t\big)}. \qquad (34)$$

For the same x^r, obviously, I have

$$EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^{0}, b^{0}; x^{r})\; EPI^{in}(g^{\mathrm{s}}, b^{\mathrm{s}}, g^{0}, b^{0}; x^{r})$$

$$= \frac{D_i(\, g^{0}; x^{r}|\Im_b^t)}{D_i(g^{t}; x^{r}|\Im_b^t)} / \frac{D_i\big(\, b^{0}; x^{r}|\Im_g^t\big)}{D_i\big(b^{t}; x^{r}|\Im_g^t\big)} = EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^{0}, b^{0}; x^{r}).$$

Q.E.D

Proposition 3.8 (Dimensionality) Suppose that the technology sets $\Im_g^t$ and $\Im_b^t$ show constant returns to scale. For a given reference vector x^r, I have $EPI^{in}(\theta g^{\mathrm{t}}, \lambda b^{\mathrm{t}}, \theta g^{0}, \lambda b^{0}; x^{r}) = EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^{0}, b^{0}; x^{r})$ for $\theta > 0$ and $\lambda > 0$.

Proof:

From Corollary to Proposition 3.5b, I have

$$EPI^{in}(\theta g^{\mathrm{t}}, \lambda b^{\mathrm{t}}, \theta g^{0}, \lambda b^{0}; x^{r}) = \frac{\theta\lambda}{\lambda\theta} EPI^{in}(g^{\mathrm{t}}, b^{\mathrm{t}}, g^{0}, b^{0}; x^{r}).$$

Q.E.D

I mention an ideal property of the EPI in Section 2, namely that the index is independent of the reference vectors. This means that for the change from (g^{t-1}, b^{t-1}) to (g^{t}, b^{t}), the value of the input-oriented EPI remains the same for any arbitrary reference vector x^r. An ideal property of this index is that the EPI scores should not change under different reference vectors. In other words, a good environmental performance index should be independent of the reference vectors. The relevant condition is on the input side, as shown in the following proposition.

Proposition 3.9 (Independent) Suppose that the technology sets $\Im_g^t$ and $\Im_b^t$ show constant returns to scale. Consider the desirable output vectors g^{t-1} and g^t and the undesirable output vectors b^{t-1} and b^t. For any reference vector x_φ^r and x_ω^r,

$$EPI^{in}\left(g^{t-1}, b^{t-1};\ g^{t}, b^{t}; x_\varphi^r\right) = EPI^{in}(g^{t-1}, b^{t-1}; g^{t}, b^{t}; x_\omega^r) \qquad (35)$$

if and only if

$$D_i(x;\ g) = w(x)D_i(\mathbf{1_N};\ g) \qquad (36)$$

and

$$D_i(x;\ b) = w(x)D_i(\mathbf{1_N};\ b). \qquad (37)$$

Proof:

Suppose that Equations (36) and (37) hold,

$$EPI^{in}\left(g^{t-1}, b^{t-1};\ g^{t}, b^{t}; x_\varphi^r\right)$$

$$= \frac{w\left(x_\varphi^r\right)D_i(\mathbf{1_N};\ g^{t-1}|\Im_b^t)}{w\left(x_\varphi^r\right)D_i(\mathbf{1_N};\ g^{t}|\Im_b^t)} \Big/ \frac{w\left(x_\varphi^r\right)D_i\left(\mathbf{1_N};\ b^{t-1}|\Im_g^t\right)}{w\left(x_\varphi^r\right)D_i\left(\mathbf{1_N};\ b^{t}|\Im_g^t\right)}$$

$$= \frac{D_i(\mathbf{1_N};\ g^{t-1}|\Im_b^t)D_i\left(\mathbf{1_N};\ b^{t}|\Im_g^t\right)}{D_i(\mathbf{1_N};\ g^{t}|\Im_b^t)D_i\left(\mathbf{1_N};\ b^{t-1}|\Im_g^t\right)}.$$

As the reference vector x_φ^r does not appear on the right side, the value of EPI^{in} for $(g^{t-1}, b^{t-1}; g^{t}, b^{t})$ will be the same, regardless of the reference vectors. This means that

$$EPI^{in}\left(g^{t-1}, b^{t-1};\ g^{t}, b^{t}; x_\varphi^r\right) = EPI^{in}(g^{t-1}, b^{t-1};\ g^{t}, b^{t}; x_\omega^r).$$

Conversely, I have

$$\frac{Q_g^{in}\left(g^{t-1}, g^t; x_\varphi^r|\Im_b^t\right)}{Q_b^{in}\left(b^{t-1}, b^t; x_\varphi^r|\Im_b^t\right)} = \frac{Q_g^{in}(g^{t-1}, g^t; x_\omega^r|\Im_b^t)}{Q_b^{in}\left(b^{t-1}, b^t; x_\omega^r|\Im_g^t\right)}. \qquad (38)$$

Given that (g^{t-1}, g^t) do not appear in the denominators and (b^{t-1}, b^t) do not appear in the numerators, it is obvious that

$$Q_g^{in}\left(g^{t-1}, g^t; x_\varphi^r | \Im_b^t\right) = \gamma \cdot Q_g^{in}(g^{t-1}, g^t; x_\omega^r | \Im_b^t). \tag{39}$$

$$Q_b^{in}\left(b^{t-1}, b^t; x_\varphi^r | \Im_g^t\right) = \gamma \cdot Q_b^{in}\left(b^{t-1}, b^t; x_\omega^r | \Im_g^t\right). \tag{40}$$

where $\gamma > 0$.

Equations (39) and (40) hold for any x_φ^r if and only if (36) and (37) are true.

Q.E.D

There is a difference between Proposition 3.2 and Proposition 3.9. Consider the following definition.

Definition: The production technology $\Im^t$ is *input-homothetic* if the input distance function is separable from the input vector in the form of

$$D_i(x; g, b) \coloneqq \sup_\lambda \left\{\lambda \colon \left(\frac{x}{\lambda}; g, b\right) \in \Im^t\right\} = w(x) D_i(1_N; g, b),$$

where $w(x)$ is a linearly homogeneous and weakly increasing function. Input homotheticity is a sufficient condition for the second part of Proposition 7, as shown in the following lemma.

Corollary to Proposition 3.9 If the production technology is input-homothetic, then

$$D_i(x;\ g | \Im_b^t) = w(x) D_i(\mathbf{1}_N;\ g | \Im_b^t) \tag{41}$$

and

$$D_i\left(x;\ b | \Im_g^t\right) = w(x) D_i\left(\mathbf{1}_N;\ b | \Im_g^t\right). \tag{42}$$

Proof:

Based on the definition of the input-oriented distance function (22), it is obvious that $(x; g, b) \in \Im^t$ if and only if $D_i(x; g, b) \geqq 1$. Then,

$$D_i(x;\ g|\Im_b^t) = \sup_{\lambda} \left\{\lambda : D_i\left(\frac{x}{\lambda}; g, b\right) \geqq 1, b \geqq 0\right\}. \tag{43}$$

As the production technology is input-homothetic, $D_i(x; g, b) = w(x)D_i(1_N; g, b)$. Hence,

$$D_i(x;\ g|\Im_b^t) = \sup_{\lambda} \left\{\lambda : w\left(\frac{x}{\lambda}\right) D_i(\mathbf{1_N}; g, b) \geqq 1, b \geqq 0\right\}. \tag{44}$$

As $w(x)$ is homogeneous of degree 1,

$$D_i(x;\ g|\Im_b^t) = \sup_{\lambda} \left\{\lambda : \tfrac{1}{\lambda} w(x) D_i(\mathbf{1_N}; g, b) \geqq 1, b \geqq 0\right\}$$

$$= w(x) \sup_{\lambda / w(x)} \left\{\frac{\lambda}{w(x)} : D_i(\mathbf{1_N}; g, b) \geqq \frac{\lambda}{w(x)}, b \geqq 0\right\}$$

$$= w(x) D_i(\mathbf{1_N};\ g|\Im_b^t).$$

The proof for $D_i(x; b)$ is similar.

Q.E.D.

Thus, homotheticity in desirable and undesirable outputs is required in Proposition 3.2, whereas only input homotheticity is needed in Proposition 3.9. Although input homotheticity is still required in Proposition 3.9, it is an improvement in the sense that I have one less reference vector and no joint homotheticity is required. I expect that the input-oriented EPI scores with different reference vectors are not comparable with each other. In this case, choosing an appropriate reference vector is crucial for measuring the input-oriented EPI.

3.3 Malmquist Environmental Performance Index

The goal of environmental performance is to maximise the ratio of desirable outputs to undesirable outputs. Thus, more desirable outputs with the same amount of undesirable outputs are preferred. Färe et al. (2004) point out that the environmental performance index is similar to the productivity index. When the input reference vector is fixed, the environmental performance indices are similar to the productivity indices, using the desirable outputs as outputs and the undesirable outputs as inputs. In addition, some empirical studies treat undesirable outputs as inputs (Hu et.al., 2012; Seiford and Zhu, 2002). Environmental performance has the same rationale of productivity. Indeed, being more environmentally friendly means that the firm produces more desirable outputs and fewer undesirable outputs. The EPI is constructed from the desirable output quantity index Q_g over the undesirable output quantity index Q_b as follows:

$$EPI = \frac{Q_g}{Q_b}. \tag{45}$$

For the case with one desirable output g and one undesirable output b, the EPI is $\frac{g}{b}$. The productivity index (*PI*) is constructed from the output quantity index Q_y over the input quantity index Q_x as follows:

$$PI = \frac{Q_y}{Q_x}. \tag{46}$$

For the case with one output y and one input x, PI is $\frac{y}{x}$. Both desirable outputs and normal outputs are good outputs that I want to maximise in the production process. Conversely, undesirable outputs and inputs are expensive and unavoidable and I want to minimise them in the production process. For environmental performance, I can consider both the benefit and the cost of the social aspect. Suppose that the total social Welfare $W = \alpha G - \mu B$ is defined as the benefit αG from the desirable output vector G at the output value vector α minus the Welfare loss μB from the undesirable output vector by-product B at the cost vector μ. Environmental performance is defined as $\frac{G}{B}$ when α and μ are fixed. Therefore, a higher EPI score shows a higher level of social Welfare with more benefits from desirable outputs or lower cost from undesirable outputs. For the productivity index showing productivity performance, I consider the economic profit of the producer. Suppose that the profit $\pi = py - cx$ is defined as the revenue py from the output vector y at the price vector p minus the cost cx from the input vector x at the cost vector c. Productivity is defined as $\frac{y}{x}$ when the prices of inputs and outputs are fixed. Therefore, higher productivity shows higher profits with more revenue from outputs or lower cost from inputs. As I treat the desirable outputs as normal outputs and the undesirable outputs as inputs, the EPI can be treated as the productivity index. Following this reasoning, I can construct the EPI from productivity index.

3.3.1 Malmquist EPI with Farrell-type distance function

As the Malmquist index is one of the most popular productivity indices, I use it to construct the environmental performance index. Consider the Malmquist productivity index (MPI) that captures productivity using input-oriented distance functions. The Malmquist productivity index uses the pair of inputs and outputs to capture productivity. For periods t and $t-1$ and the productivity indices for the pair of inputs and outputs for each period (x^{t-1}, y^{t-1}) and (x^t, y^t), the input-oriented Malmquist productivity index for $\mathfrak{J}^t$ is,

$$MPI_i^t = \frac{D_i^t(x^t, y^t)}{D_i^t(x^{t-1}, y^{t-1})}. \tag{47}$$

Similarly, the output-oriented Malmquist productivity index for $\mathfrak{J}^t$ is,

$$MPI_o^t = \frac{D_o^t(x^t, y^t)}{D_o^t(x^{t-1}, y^{t-1})}. \tag{48}$$

For a pair of desirable outputs and undesirable outputs (g^{t-1}, b^{t-1}) and (g^t, b^t), I set the input reference vector as x^r. Thus, the change of distance functions from $D(g^{t-1}, b^{t-1}; x^r)$ to $D(g^t, b^t; x^r)$ shows the change of environmental performance from (g^{t-1}, b^{t-1}) to (g^t, b^t). The *Malmquist Environmental Performance Index* (MEPI) with desirable output-oriented distance function for $\mathfrak{J}^t$ is

$$MEPI_g^t(g^t, b^t; g^{t-1}, b^{t-1}; x^r) = \frac{D_g(g^t, b^t; x^r|\mathfrak{I}^t)}{D_g(g^{t-1}, b^{t-1}; x^r|\mathfrak{I}^t)}. \tag{49}$$

Similarly, the *Malmquist environmental performance index* (MEPI) with undesirable output-oriented distance function for $\mathfrak{I}^t$ is,

$$MEPI_b^t(g^t, b^t; g^{t-1}, b^{t-1}; x^r) = \frac{D_b(g^t, b^t; x^r|\mathfrak{I}^t)}{D_b(g^{t-1}, b^{t-1}; x^r|\mathfrak{I}^t)}. \tag{50}$$

Environmental performance improves when the Malmquist EPI score increases. When the desirable output vector and the undesirable output vector are identical, $(g^t, b^t) = (g^{t-1}, b^{t-1})$, the MEPI is equal to 1. When the MEPI is greater than 1, environmental performance is better in (g^t, b^t) than in (g^{t-1}, b^{t-1}). When the MEPI is less than 1, environmental performance is worse in (g^t, b^t) than in (g^{t-1}, b^{t-1}). The Malmquist EPI constructed from the Farrell-type distance function may well-defined. Consider Proposition 3.1 in Section 3.1.

In Proposition 3.1a, $D_g(g^t, b^t; x^r|\mathfrak{I}^t)$ is well-defined if and only if there exists g, such that $(x^r, g, b^t) \in \mathfrak{I}^t$.And $D_g(g^{t-1}, b^{t-1}; x^r|\mathfrak{I}^t)$ is well-defined if and only if there exist g, such that $(x^r, g, b^{t-1}) \in \mathfrak{I}^t$. $D_b\ (g^t, b^t; x^r|\mathfrak{I}^t)$ is Well defined if and only if $b^t \in K(x^r, g^t)$ and $D_b\ (g^{t-1}, b^{t-1}; x^r|\mathfrak{I}^t)$ is well-defined if and only if $b^{t-1} \in K(x^r, g^{t-1})$.

Corollary to Proposition 3.1a $MEPI_g^t(g^t, b^t; g^{t-1}, b^{t-1}; x^r)$ is a finite number if and only if there exist g, such that $(x^r, g, b^t) \in \mathfrak{I}^t$and $(x^r, g, b^{t-1}) \in \mathfrak{I}^t$.

Corollary to Proposition 3.1b $MEPI_b^t(g^t, b^t; g^{t-1}, b^{t-1}; x^r)$ is a finite number if and only if $b^t \in K(x^r, g^t)$ and $b^{t-1} \in K(x^r, g^{t-1})$.

These two corollaries are the same as the corollary to Proposition 3.1. As previous example in Appendix 3.4, EPI^{out}may have no solution. Similarly, these corollaries indicate that the MEPI with Farrell-type distance function is not feasible for some reference vectors and frontier. I propose another version of the EPI to solve this infeasibility problem.

3.3.2 Malmquist EPI with unrestricted input production set

In the input-oriented EPI, I use the unrestricted production set with desirable outputs and undesirable outputs to ensure the feasibility of the input-oriented distance function. Similarly, I construct the unrestricted input production set to solve the infeasibility problem of the output-oriented distance function. I use the output-oriented distance function with unrestricted inputs in the output-oriented EPI. Let's define the *production set with unrestricted inputs* as $\Im_i^t \coloneqq \{(g,b): (x; g, b) \in \Im^t, x \geqq 0\}$. In this production set, (g,b) are feasible as long as they can be produced for some $x \geqq 0$. The *desirable output-oriented distance function* for (g,b) with respect to the production set $\Im_i^t$ is

$$D_g(g, b|\Im_i^t) \coloneqq \inf_\theta \left\{\theta: \left(\frac{g}{\theta};\ x, b\right) \in \Im^t, x \geqq 0\right\}. \tag{51}$$

Similarly, the *undesirable output-oriented distance function* for (g, b) with respect to the production set $\mathfrak{I}_i^t$, is

$$D_b(g; b|\mathfrak{I}_i^t) \coloneqq \sup_{\theta}\left\{\theta\colon \left(\frac{b}{\theta};\ x, g\right) \in \mathfrak{I}^t, x \geqq 0\right\}. \tag{52}$$

From the desirable output-oriented distance function and the undesirable output-oriented distance function, I construct the MEPI with respect to the unrestricted input production set $\mathfrak{I}_i^t$ as

$$MEPI_g^t(g^t, b^t; g^{t-1}, b^{t-1}) = \frac{D_g(g^t, b^t|\mathfrak{I}_i^t)}{D_g(g^{t-1}, b^{t-1}|\mathfrak{I}_i^t)}. \tag{53}$$

$$MEPI_b^t(g^t, b^t; g^{t-1}, b^{t-1}) = \frac{D_b(g^t, b^t|\mathfrak{I}_i^t)}{D_b(g^{t-1}, b^{t-1}|\mathfrak{I}_i^t)}. \tag{54}$$

Similar to the MEPI using the Farrell-type distance function, the unrestricted Malmquist EPI shows the change in environmental performance. Environmental performance improves when the EPI score increases. When the desirable output vector and the undesirable output vector are identical, $(g^t, b^t) = (g^{t-1}, b^{t-1})$, the MEPI is equal to 1. When the MEPI is greater than 1, environmental performance is better in (g^t, b^t) than in (g^{t-1}, b^{t-1}). When the MEPI is less than 1, environmental performance is worse in (g^t, b^t) than in (g^{t-1}, b^{t-1}). Proposition 3.10 checks the feasible conditions of the undesirable output-oriented MEPI and the desirable output-oriented MEPI.

Proposition 3.10 Let $\mathfrak{I}$ be a production set with constant returns to scale. Suppose that inputs and desirable outputs are strongly disposable, while undesirable outputs are Weakly disposable. Suppose that $D_g(g, b|\mathfrak{I}_i^t)$ is a finite number for $(g, b) \in \Im_i^t$. Consider the reference vectors $g^r \in \mathbb{R}_{++}^M$ and $b^r \in \mathbb{R}_{++}^L$. Let $(g^t, b^t) \in \Im_i^t$.

a. If $(g, b^r) \in \Im_i^t$ for some $g > 0$, then $D_g(g^t, b^r|\Im_i^t)$ is a finite number.

b. Define a cone: $K(g^r) \coloneqq \{b > 0 \colon \alpha b \in P^t(g^r)\ for\ some\ \alpha > 0\}$. Then $D_b(g^r, b^t|\Im_i^t)$ is a finite number if and only if $b^t \in K(g^r)$.

For sequential frontiers, it is always true that $(g^t, b^t) \in \Im_i^t$ and $(g^{t-1}, b^{t-1}) \in \Im_i^t$. Thus, $MEPI_g^t(g^t, b^t; g^{t-1}, b^{t-1}|\Im_i^t)$ is well-defined from Proposition 3.10a. In addition, as $b^t \in K(g^t|\Im_i^t)$ and $b^{t-1} \in K(g^{t-1}|\Im_i^t)$, $MEPI_b^t(g^t, b^t; g^{t-1}, b^{t-1}|\Im_i^t)$ is well-defined from Proposition 3.10b. Therefore, the MEPIs with respect to the unrestricted inputs production set can achieve feasible solutions with sequential frontiers and prior-lag reference vectors. Similar to the desirable properties of the input-oriented EPI, the MEPI (both $MEPI_g$ and $MEPI_b$) satisfies the following properties.

Homogeneity:

$$MEPI(\alpha g^t, \beta b^t, \gamma g^{t-1}, \delta b^{t-1}|\Im_i^t) = \frac{\alpha\delta}{\beta\gamma} MEPI(g^t, b^t, g^{t-1}, b^{t-1}|\Im_i^t)$$

for $\theta > 0$, $\lambda > 0$, $\alpha > 0$, and $\beta > 0$.

Dimensionality:

$$MEPI(\theta g^t, \lambda b^t, \theta g^{t-1}, \lambda b^{t-1}|\Im_i^t) = MEPI(g^t, b^t, g^{t-1}, b^{t-1}|\Im_i^t) \text{ for } \theta > 0$$

and $\lambda > 0$.

Proposition 3.11 (Homogeneity) Suppose that the production set with unrestricted inputs $\Im_i^t$ shows constant returns to scale. Suppose that the previous technology is contained in the current technology $\Im_i^{t-1} \subset \Im_i^t,\ t = 2, \cdots, T$. In addition, desirable outputs and undesirable outputs are jointly weakly disposable. For $\alpha, \beta, \gamma, \delta > 0$,

$$MEPI_g(\alpha g^t, \beta b^t; \gamma g^{t-1}, \delta b^{t-1} | \Im_i^t) = \frac{\alpha\delta}{\beta\gamma} MEPI_g(g^t, b^t; g^{t-1}, b^{t-1} | \Im_i^t).$$

From the property of homogeneity, I can directly derive dimensionality. Dimensionality is an important property for a good index that is independent of the measure of the unit. This can be shown as follows.

Corollary to Proposition 3.11 (Dimensionality) Suppose that the production set with unrestricted inputs $\Im_i^t$ shows constant returns to scale. Suppose that the previous technology is contained in the current technology $\Im_i^{t-1} \subset \Im_i^t,\ t = 2, \cdots, T$. In addition, desirable outputs and undesirable outputs are jointly Weakly disposable. For $\theta > 0$ and $\lambda > 0$,

$$MEPI_g(\theta g^t, \lambda b^t;\ \theta g^{t-1}, \lambda b^{t-1} | \Im_i^t) = MEPI_g(g^t, b^t, g^{t-1}, b^{t-1} | \Im_i^t).$$

Both the Malmquist EPI and the input-oriented EPI are finite numbers for various benchmark production frontiers and reference vectors. They are superior to the output-oriented EPI in this sense. The only difference between these two measures is that the Malmquist EPI is constructed from the productivity index, whereas the input-oriented EPI is constructed from the quantity index.

3.3.3 Decomposition of the Malmquist EPI

The formula of the Malmquist productivity index is used to construct the Malmquist EPI in the previous subsection. As the Malmquist productivity index can be decomposed into technological change and efficiency change, I can decompose the change in environmental performance into environmental technological change and environmental efficiency change. Consider the decomposition of the output-oriented Malmquist productivity index as follows (the input-oriented Malmquist productivity index being similar, I skip this discussion)

$$MPI_o^t = \frac{D_o^t(x^t, y^t)}{D_o^t(x^{t-1}, y^{t-1})} = \underset{\text{efficiency change}}{\frac{D_o^t(x^t, y^t)}{D_o^{t-1}(x^{t-1}, y^{t-1})}} \underset{\text{technological change}}{\frac{D_o^{t-1}(x^{t-1}, y^{t-1})}{D_o^t(x^{t-1}, y^{t-1})}}.$$

Similarly, the Malmquist productivity index with unrestricted inputs of the desirable output-oriented approach and the undesirable output-oriented approach can be broken down into

$$MEPI_g^t = \frac{D_g^t(g^t, b^t)}{D_g^t(g^{t-1}, b^{t-1})} = \underset{\text{environmental efficiency change}}{\frac{D_g^t(g^t, b^t)}{D_g^{t-1}(g^{t-1}, b^{t-1})}} \underset{\text{environmental technological change}}{\frac{D_g^{t-1}(g^{t-1}, b^{t-1})}{D_g^{t-1}(g^{t-1}, b^{t-1})}}.$$

$$MEPI_b^t = \frac{D_b^t(g^t, b^t)}{D_b^t(g^{t-1}, b^{t-1})} = \underset{\text{environmental efficiency change}}{\frac{D_b^t(g^t, b^t)}{D_b^{t-1}(g^{t-1}, b^{t-1})}} \underset{\text{environmental technological change}}{\frac{D_b^{t-1}(g^{t-1}, b^{t-1})}{D_b^{t-1}(g^{t-1}, b^{t-1})}}.$$

Hence, when there is a change in environmental performance indicated by the MEPI, the above decomposition helps decision makers to identify the source of change: Does it come from catching-up effect (moving towards the frontier), or technological change (shift of production frontier), or both? This is a definite advantage of the MEPI over the input-oriented and output-oriented EPI.

3.4 Reference Vectors and Frontiers in DEA

In the productive performance analysis of the data envelopment analysis approach, there are three commonly used reference frontiers: the current frontier, the sequential frontier and the grand frontier. The current frontier at time t $\Im^{tc}(x^t; g^t, b^t)$ is constructed from all observed productions in time t. The empirical frontier for period *t* is constructed as follows,

$$\Im^{tc} := \left\{ \begin{array}{c} (x^t; g^t, b^t) \colon \sum_{k=1}^{K} z_k^t g_k^t \geqq g^t, \\ \sum_{k=1}^{K} z_k^t b_k^t = b^t, \\ \sum_{k=1}^{K} (z_k^t + \mu_k^t) x_k^t \leqq x^t \ , \\ \mu_t^k \geqq 0, z_t^k \geqq 0 \ for \ k = 1, \cdots K \end{array} \right\}. \tag{55}$$

The corresponding production set with unrestricted undesirable outputs for the current frontier $\Im_b^{tc}$ is $\Im_b^{tc} := \{(x^t, g^t) \colon (x^t; g^t, b^t) \in \Im^{tc}, b^t \geqq 0\}$. The empirical frontier is,

$$\Im_b^{tc} := \left\{ \begin{array}{c} (x^t, g^t): \sum_{k=1}^{K} z_k^t g_k^t \geqq g^t, \\ b^t \geqq 0, \\ \sum_{k=1}^{K} (z_k^t + \mu_k^t) x_k^t \leqq x^t \ , \\ \mu_t^k \geqq 0, z_t^k \geqq 0 \; for \; k = 1, \cdots K \end{array} \right\}. \qquad (56)$$

The corresponding production set with unrestricted desirable outputs for the current frontier $\Im_g^{tc}$ is $\Im_g^{tc} := \{(x^t, b^t): (x^t; g^t, b^t) \in \Im^{tc}, g^t \geqq 0\}$. The empirical frontier is,

$$\Im_g^{tc} := \left\{ \begin{array}{c} (x^t, b^t): g^t \geqq 0, \\ \sum_{k=1}^{K} z_k^t b_k^t = b^t, \\ \sum_{k=1}^{K} (z_k^t + \mu_k^t) x_k^t \leqq x^t , \\ \mu_t^k \geqq 0, z_t^k \geqq 0 \; for \; k = 1, \cdots K \end{array} \right\}. \qquad (57)$$

The sequential frontier $\Im^{ts}$ is constructed from all observed data from all previous years and the current year. When using the sequential frontier, it is assumed that there is no technological reduction, thus all previous productions must be within the current frontier. The previous technology is contained in the current technology $\Im^{t-1} \subset \Im^t, \ t = 2, \cdots, T$. The empirical frontier for period *t* is constructed as follows,

$$\Im^{ts} := \left\{ \begin{array}{c} (x^t; g^t, b^t): \sum_{i=1}^{t}\sum_{k=1}^{K} z_k^i g_k^i \geqq g^t, \\ \sum_{i=1}^{t}\sum_{k=1}^{K} z_k^i b_k^i = b^t, \\ \sum_{i=1}^{t}\sum_{k=1}^{K} \left(z_k^i + \mu_k^i\right) x_k^i \leqq x^t\,, \\ \mu_k^i \geqq 0, z_k^i \geqq 0 \\ for\ k = 1, \cdots K;\ i = 1, \cdots, t \end{array} \right\} \tag{58}$$

The corresponding production set with unrestricted undesirable outputs for the sequential frontier $\Im_b^{ts}$ is $\Im_b^{ts} := \{(x^t, g^t): (x^t; g^t, b^t) \in \Im^{ts}, b^t \geqq 0\}$. The empirical frontier is,

$$\Im_b^{ts} := \left\{ \begin{array}{c} (x^t, g^t): \sum_{i=1}^{t}\sum_{k=1}^{K} z_k^i g_k^i \geqq g^t, \\ b^t \geqq 0, \\ \sum_{i=1}^{t}\sum_{k=1}^{K} \left(z_k^i + \mu_k^i\right) x_k^i \leqq x^t, \\ \mu_k^i \geqq 0, z_k^i \geqq 0 \\ for\ k = 1, \cdots K;\ i = 1, \cdots, t \end{array} \right\}. \tag{59}$$

The corresponding production set with unrestricted desirable outputs for the sequential frontier $\Im_g^{ts}$ is $\Im_g^{ts} := \{(x^t, b^t): (x^t; g^t, b^t) \in \Im^{ts}, g^t \geqq 0\}$. The empirical frontier is,

$$\mathfrak{I}_g^{ts} \coloneqq \left\{ \begin{array}{c} (x^t, b^t) \colon g^t \geqq 0, \\ \sum_{i=1}^{t} \sum_{k=1}^{K} z_k^i b_k^i = b^t, \\ \sum_{i=1}^{t} \sum_{k=1}^{K} \left(z_k^i + \mu_k^i\right) x_k^i \leqq x^t\,, \\ \mu_k^i \geqq 0, z_k^i \geqq 0 \\ for\; k = 1, \cdots K;\; i = 1, \cdots, t \end{array} \right\}. \tag{60}$$

The last one is the grand frontier $\mathfrak{I}^*$ for the entire study period, i.e., $\mathfrak{I}^t = \mathfrak{I}^*$ such that $(x_k^t, g_k^t, b_k^t) \in \mathfrak{I}^*$ for all $k = 1, \dots, K$ and $t = 1, \dots, T$. The empirical frontier is,

$$\mathfrak{I}^* \coloneqq \left\{ \begin{array}{c} (x; g, b) \colon \sum_{t=1}^{T} \sum_{k=1}^{K} z_k^t g_k^t \geqq g, \\ \sum_{t=1}^{T} \sum_{k=1}^{K} z_k^t b_k^t = b, \\ \sum_{t=1}^{T} \sum_{k=1}^{K} (z_k^t + \mu_k^t) x_k^t \leqq x\,, \\ \mu_k^t \geqq 0, z_k^t \geqq 0 \\ for\; k = 1, \cdots K;\; t = 1, \cdots, T \end{array} \right\}. \tag{61}$$

The corresponding production set with unrestricted undesirable outputs for the grand frontier $\mathfrak{I}_b^*$ is $\mathfrak{I}_b^* \coloneqq \{(x, g) \colon (x; g, b) \in \mathfrak{I}^*, b \geqq 0\}$. The empirical frontier is,

$$\Im_b^* \coloneqq \left\{ \begin{array}{c} (x,g)\colon \sum_{t=1}^{T}\sum_{k=1}^{K} z_k^t g_k^t \geqq g, \\ b \geqq 0, \\ \sum_{t=1}^{T}\sum_{k=1}^{K} (z_k^t + \mu_k^t) x_k^t \leqq x\,, \\ \mu_k^t \geqq 0, z_k^t \geqq 0 \\ for\ k = 1,\cdots K;\ t = 1,\cdots,T \end{array} \right\}. \qquad (62)$$

The corresponding production set with unrestricted desirable outputs for the grand frontier $\Im_g^*$ is $\Im_g^* \coloneqq \{(g,b)\colon (x;g,b) \in \Im^*, g \geqq 0\}$. The empirical frontier is

$$\Im_g^* \coloneqq \left\{ \begin{array}{c} (x,b)\colon g \geqq 0, \\ \sum_{t=1}^{T}\sum_{k=1}^{K} z_k^t b_k^t = b, \\ \sum_{t=1}^{T}\sum_{k=1}^{K} (z_k^t + \mu_k^t) x_k^t \leqq x\,, \\ \mu_k^t \geqq 0, z_k^t \geqq 0 \\ for\ k = 1,\cdots K;\ t = 1,\cdots,T \end{array} \right\}. \qquad (63)$$

One of the disadvantages of the grand frontier is that the technology is fixed and does not show dynamic changes. It will overestimate the production set for the first serval years. In addition, the grand frontier will be changed for all years when new data are added. The most appropriate frontier for capturing dynamic change is the current frontier, which can capture technological change based on current observed data. In the analysis of productivity and efficiency, the current frontier is the common reference frontier. Unfortunately, current frontier may cause infeasibility problem in EPI. Thus, I suggest sequential frontier as reference frontier in input-oriented EPI and Malmquist EPI.

Next, I discuss the reference vectors to be used in the EPI evaluation. Currently, two reference vectors are used. The first vector is the min-max reference vector used in the output-oriented EPI. This reference vector is constructed from the minimum desirable output vector and the maximum input vector. The min-max reference vector in time *t* is,

$$R_t{}^{min-max} \coloneqq \left(x^{max}, g^{min}\right)$$

Where $x^{max} = (x_1^{max}, \cdots, x_N^{max})$ and $g^{min} = \left(g_1^{min}, \cdots, g_M^{min}\right)$,

$$x_n^{max} := \max_{k,t}\{x_{kn}^t : k = 1, \cdots, K; t = 1, \cdots, T\}, n = 1, \cdots, N.$$

$$g_m^{min} := \min_{k,t}\{b_{km}^t : k = 1, \cdots, K; t = 1, \cdots, T\}, m = 1, \cdots, M.$$

The problem is that there is no explicit economic interpretation for the min-max reference vectors. In addition, the min-max reference vectors are constructed from extreme values, which are sensitive to outliers. The values of output-oriented EPI computed based on the min-max reference vectors will be sensitive to outliers, too. It seems the only reason of using these reference vectors is to ensure feasible solutions. When the EPI formula allows for flexible choices of reference vectors, the min-max reference vectors can be abandoned.

Another commonly used reference vector in the analysis of productivity performance uses the observations of the previous year, which I call the lag reference vectors R^{lag}. The EPI uses these reference vectors to reflect the dynamic change in the same decision-making unit between last year and the current year. The lag reference vectors $R^{lag}(x^{t-1}; g^{t-1}, b^{t-1})$ in period *t* for firm *k* are,

$$R_t^{lag} \coloneqq (x_k^{t-1}; g_k^{t-1}, b_k^{t-1}).$$

These reference vectors are commonly used in productivity indices, such as the Malmquist index. The EPI will start from the second year (the EPI score for the first year is used as benchmark). If the EPI score is greater than 1, the environmental performance of this decision-making unit is better than that of the previous year. Conversely, if the EPI score is less than 1, environmental performance is worse than the previous year. There are some limitations. First, there is no cross-sectional comparison between different decision-making units. I can only evaluate the improvement in environmental performance for each unit. Second, I cannot evaluate environmental performance in the first year. Finally, the

EPI score reflects the change in environmental performance, not the level of environmental performance.

An alternative dynamic reference vector is the current reference vector$s$ $R^{cur}(x^t; g^t, b^t)$, which use next year's observations as the reference. The formula of $R^{cur}(x^t; g^t, b^t)$ in period t for firm k is

$$R_t^{cur} \coloneqq (x_k^t; g_k^t, b_k^t).$$

I can take the geometric mean of the EPI score from the 'prior' reference vectors and the lag reference vectors. I call this approach the prior-lag reference vectors. Next, one may consider using the average of the current observations as reference vectors, which I call the current-average reference vectors R^{ca}. The EPI computed from these reference vectors shows the relative environmental performance compared with the current average level. The formula of R^{ca} in period *t* is

$$R_t^{ca} \coloneqq \left(\frac{\sum_{k=1}^{K} x_t^k}{K}; \frac{\sum_{k=1}^{K} g_t^k}{K}, \frac{\sum_{k=1}^{K} b_t^k}{K}\right).$$

For the current average reference vectors, if the EPI score is greater than 1, environmental performance is better than the average level of environmental performance that year. If one concerns the environmental performance of a unit relative to the average, then the current average reference vectors are a good choice. If the objective is to monitor the environmental performance of the same unit over time, then other reference vectors should be chosen.

Finally, I propose a fixed reference vector that can compare the change in environmental performance with respect to in initial situation. I suggest the first-year average reference vectors R^{fa}, which are constructed from the averages of the inputs and outputs observed the first year of the data set. The reference vector is the same for all time *t*,

$$R^{fa} := \left(\frac{\sum_{k=1}^{K} x_1^k}{K}; \frac{\sum_{k=1}^{K} g_1^k}{K}, \frac{\sum_{k=1}^{K} b_1^k}{K}\right).$$

Using this formula, the EPI score reflects the relative environmental performance compared with the environmental performance of the first-year average. If the EPI score is greater than 1, environmental performance is better than the environmental performance of the average production in the initial year. Thus, the environmental performance of all observations of the entire period can be compared.

For empirical studies, the current frontier with constant returns to scale is the best choice. For the choice of reference vectors, there are two better choices with regard to the research question. First, for a cross-sectional comparison, to investigate the difference between different production units, the fixed reference vectors with current year average are the best choice. Second, for a dynamic comparison, the current-lag reference vectors are the best as they compare the change in the same production unit between last year and the current year.

3.5 Conclusions

The output-oriented environmental performance index is not well-defined without restrictive conditions, namely the grand frontier and the min-max reference vectors. Therefore, I propose alternative approaches to construct the environmental performance index. Two approaches are introduced in this chapter. First, the environmental performance index is the ratio between the desirable output quantity index and the undesirable output quantity index.

My contribution is to replace the output-oriented distance functions in the formulas of output quantity index by input-oriented distance functions. Second, the environmental performance index is the ability to produce desirable outputs against undesirable outputs as by-products. Thus, I use the Malmquist productivity index to construct the Malmquist EPI. In both approaches, the newly introduced concepts of unrestricted production sets are crucial. The two modified EPIs improve upon the existing output-oriented EPI by ensuring the existence of solutions with more benchmark production frontiers and reference vectors. An additional contribution of this chapter is the ability of the Malmquist EPI to decompose the change in environmental performance into technological change and efficiency change.

Compared with other approaches, the MEPI can identify the source of environmental performance change: moving towards the frontier or shift of production frontier. In addition to the decomposition. It uses the pair of desirable outputs and undesirable outputs of the current observations. To evaluate the dynamic change in environmental performance from the pair of desirable outputs and undesirable outputs, the Malmquist EPI reflects environmental change when desirable outputs and undesirable outputs are paired for a dynamic comparison.

For the choice of reference vectors, there are two better choices with regard to the research question. First, for a cross-sectional comparison, the fixed reference vectors with current-year average are the best choice. Second, for a dynamic comparison, the current-lag reference vectors are the best choice. The problem of infeasible solutions has not been completely solved. Constructing the production frontier from current years may still have no solutions for some observations. This is a direction for future research.

Chapter 4

Treatment Quality Performance in the Hospital Sector in China: Effects of Marketisation and Government Subsidies

The Chinese government has reformed the hospital sector over time. There are two alternative policy directions for healthcare reforms. The first policy direction is market-oriented reform, which includes four main strategies to increase the degree of marketisation in the hospital sector (Helderman et al., 2005; Harrison and Calltorp, 2000; Schlesinger, 2002). The first strategy is expanding the market share of private hospitals or for-profit hospitals in the hospital sector. The second is releasing the price regulation of medical services and reducing direct subsidies to public hospitals. The third is allowing the transformation of the ownership of public hospitals into private hospitals. The fourth is developing the medical insurance market and let people choose. These policy strategies to implement the market-oriented reform are called market-oriented strategies. The second policy direction is the government-oriented reform, which is the opposite of the market-oriented reform. The target of the government-oriented reform is to increase the degree of government intervention in the hospital sector. The policy strategies to implement the government-oriented reform are called government-oriented strategies. Government-oriented strategies are in the opposite direction of market-oriented strategies. Key points of the government-oriented strategies are as follows (He and Tian, 2008; Ho, 2010; Bai et al., 2006). First, expanding the market share of public hospitals or non-profit hospitals in the hospital sector.

Second, increasing the degree of price regulation of medical services and increasing direct subsidies to public hospitals. Third, encouraging the transformation of the ownership of private hospitals into public hospitals. Finally, creating a compulsory social health insurance system for citizens.

Health care reforms have shifted back and forth between market-oriented and government-oriented policies (Blomqvist, 2009). China's healthcare reform began with the market-oriented reform, which implemented administrative decentralisation and then added market forces to the hospital sector (Akin et al., 2005). Before the 1980s, all medical organisations in China were state-owned. However, with the rapid growth of the Chinese economy, both medical demand and medical costs increased drastically. To solve the financial problems of the healthcare sector, China launched the market-oriented reform in 1984, which continued until 2005. On the one hand, the ownership reform encouraged the establishment of new private hospitals and the government sold some state-owned hospitals to private investors. On the other hand, state-owned hospitals became more financially independent with the financial reform. In 2005, the central government announced the failure of the previous market-oriented healthcare reform[2]. In contrast to the market-oriented reform, the government then limited market competition by increasing financial support to public hospitals and controlling the price of medical services at a low level. Since 2009, the

---

[2] The report of the healthcare reform, in Chinese

government has tried to strike a balance between market-oriented policies and government invention, increasing subsidies and encouraging the creation of more private hospitals to increase competition.

Market-oriented reform can have positive and negative effects on the healthcare sector. On the one hand, Hu et.al. (2012) believes that market competition can promote efficiency and productivity in the healthcare sector. On the other hand, Forder and Allan (2014) argues that the profit target with asymmetric information can lead hospitals to make less effort in the quality of treatment. Many studies point out the drawbacks of market-oriented reform that it can increase the cost of medical services and reduce their quality (Forder and Allan 2014; Liu, 2005; Gu, 2005). On the other hand, government intervention also has problem on cause inefficiency and shortages of supply. (Gerald and Gu, 1997; Feng and Yu, 2008; Chen et al., 2008).

As Blomqvist (2009) points out, in the absence of government intervention, a high degree of market failure can be expected. In contrast, when the government is too involved in the provision of healthcare services, inefficient operations occur. Currently, as in other countries, China has both for-profit healthcare institutions and government-funded institutions. The net effects of marketisation and government intervention are not yet clear. Applying the modified EPIs in the last chapter, this chapter seeks to determine the relative merits of higher level of marketisation and greater government involvement from the point of view of quality performance of the hospital sector in China.

One of the main measures of market-oriented policies is to increase the number of private hospitals. This measure promotes competition in the hospital sector, leading hospitals to pay more attention to costs and revenue. This should result in more medical services using fewer resources. However, deaths are inevitable during the process of providing medical services and can be regarded as undesirable outputs in the production process. Reducing the number of deaths during treatments is also an important consideration for providing medical services from a societal perspective. In this chapter, treatment quality is defined as the amount of treatment (medical services) per death. It should be noted that pursuing less deaths during treatments may not increase profits. Hence a higher treatment quality is not necessarily consistent with profit maximisation. The net effects of increasing market competition on treatment quality are unknown.

In contrast, one of the main measures of government-oriented policies is to provide direct subsidies to hospitals. With less competitive market pressure and more subsidies, doctors can focus on treating patients and avoiding deaths simultaneously. However, less competitive pressure can reduce productive efficiency. The final consequences of more subsidies on treatment quality are also unknown. Currently, private hospitals and public hospitals coexist in China. The issue of increasing the number of private hospitals is under intense discussion. By ascertaining the effects of a higher level of marketisation and greater government involvement on treatment quality in the hospital sector, this chapter aims to provide implications for both policies.

Some studies focus on treatment quality and government subsidies in the hospital sector, such as Bloom, Propper and Van (2015) and Duggan (2000).

However, they do not consider these variables simultaneously for market-oriented and government-oriented policies. To conduct a thorough analysis of the performance of the hospital sector in China, this chapter adopts a production approach that integrates the multidimensional nature of production variables and treatment quality.

Hospitals can be treated as production units, with desirable outputs (medical services), undesirable outputs (deaths)[3] and inputs. This chapter adopts the data envelopment analysis (DEA) approach to study the performance of treatment quality of hospitals in China. The current literature using the DEA approach to investigate hospital performance focuses on technical efficiency and productivity, see Hollingsworth (2008) for a survey. By adopting the environmental performance index (EPI) discussed in Essay 2 to reflect the quality of services provided by hospitals, I am able to identify the patterns of treatment quality of Chinese hospitals during the study period and across regions. In this chapter, I call the EPI the treatment quality performance index (TQI) to evaluate the performance of treatment quality. By examining the determinants of treatment quality, I can quantify the effects of market-oriented and government-oriented policies. This can help policymakers develop better policies to reform the healthcare sector.

The rest of this chapter is organised as follows. Section 1 provides

---

[3] Undesirable outputs produced by hospitals include more items than deaths. Only deaths are considered here because various types of deaths during treatments are the most important undesirable outputs and because of availability of data.

background information on the hospital sector in China. Section 2 explains the methodology for estimating treatment quality. Section 3 describes the data and presents the estimates of treatment quality in China. Section 4 explores the determinations of TQI and policy implications. Section 5 discusses policy implications and Section 6 concludes the essay.

4.1 Background Information

Before 1978, all hospitals in China were public hospitals. For urban residents, all treatment costs were covered by their employers. For rural residents, the price of medical services was controlled by the central government at a low level. The mortality rate fell from 1.23% in 1955 to 0.62% in 1978 and life expectancy increased from 35 to 66 years during the same period. However, both medical demand and medical costs increased drastically. From 1965 to 1985, medical expenditure increased more than 28 times, while GDP increased only about 13 times. Thus, the old central planning medical system could no longer cover the huge medical expenditure.

The market-oriented reform in the Chinese hospital sector began in 1984. It had three main goals: increasing market competition with ownership reform, increasing financial independence with financial reform and reducing price regulation with price reform (Yip and Hsiao, 2009; Blumenthal and Hsiao, 2015). First, ownership reform allowed the creation of private hospitals and some public hospitals were sold to private investors. Second, public hospitals became more financially independent with financial reform. The share of public hospital revenue based on government subsidies decreased from 32.16% in 1978 to 13.24% in 2015. Finally, hospitals had the power to determine the prices of medical services and drugs, except for those listed in the basic price regulation of medical services and drugs, with price reform. These market-oriented strategies solved the financial problems of hospitals and encouraged investments in the hospital sector. However, the price of medical services increased dramatically after the market-oriented reform.

In 2005, the central government announced the failure of the previous market-oriented healthcare reform (The report of the healthcare reform, in Chinese). According to the report of the healthcare reform, the market-oriented reform caused two major problems. One was the dramatic increase in unregulated treatment costs and the other was the decrease in the quality of treatment. Furthermore, many studies reveal that the market-oriented reform led to expensive and low-quality medical services (Liu, 2005; Gu, 2005). As opposed to the market-oriented reform, the government then limited market competition by increasing financial support to public hospitals and controlling the price of medical services at a low level. However, many researchers show that government intervention reduced efficiency and increased supply shortages (Gerald and Gu, 1997; Feng and Yu, 2008; Chen et al., 2008).

After 2009, the government tried to strike a balance between marketisation and government intervention. Government subsidies Were increased. Financial support to the hospital sector increased from RMB482 billion in 2009 to RMB1,058 billion in 2014. The government also encouraged the creation of more private hospitals to increase competition. The ratio of private hospitals in the hospital sector increased rapidly from 28% in 2009 to 49% in 2014. These new policies are part of the ‘new healthcare reform’, which had three main objectives: increasing the supply of medical services in response to the shortage problem, reducing the financial burden of patients and improving treatment quality.

In this chapter, policies that focus on market forces are called *market-oriented strategies*. Conversely, policies that focus on government intervention are called *government-oriented strategies*. When market forces and government intervention are of equal importance, I call this a *mixed strategy*. Researchers do not agree on the roles of marketisation and government intervention in the hospital sector. Therefore, there is no conclusion about the optimal strategy. One view is that there is a market failure in the hospital sector and that private hospitals will increase the price of medical services and reduce investment in prevention. For example, Liu (2004) points out that as hospital financing in China became increasingly privatised, hospitals were less interested in public health work. The other view is that inefficiency and supply shortages in the hospital sector in China are due to price regulation and the vacancy of hospitals’ independent management right. Chen et al. (2008) find that due to lack of competition, excessive government subsidies led to the inefficiency of public hospitals.

O’Newell et al. (2008) conduct a systematic review of empirical studies on the hospital sector using the DEA approach published between 1984 and 2004. Moreover, in general, private hospitals are more efficient than public hospitals in

Germany and the U.S. (Herr, 2008; Ozcan et al., 1992). Previous studies also show evidence to support the market-oriented ownership reform (Sloan et al., 2001; Kolstad and Kowalski, 2012; López-Casasnovas and Puig-Junoy, 2000). However, most studies focus on technical efficiency and do not consider undesirable outputs. Hu et.al. (2012) incorporate undesirable outputs in their study and find that a high proportion of private hospitals and high-quality hospitals can help increase technical efficiency. These studies focus mainly on the technical efficiency of hospitals.

From my point of view, more private hospitals can promote competition in the hospital sector, with hospitals providing medical services using fewer resources. Regardless of the production techniques used to treat patients, this strategy can push hospitals to provide more medical services with the same amount of resources, which may improve treatment quality. However, there may be several ways to treat the same disease and the higher the cost, the lower the probability of death. Although reducing the number of deaths is also an important consideration in the provision of medical services from a societal perspective, it needs not be consistent with profit maximisation. The final consequences of increased competition on treatment quality are unknown. Conversely, with more government subsidies, the pressure of market competition is lower. Therefore, doctors can focus on treating patients and avoiding deaths simultaneously. However, less competitive pressure can reduce productive efficiency. The final consequences of more government subsidies on treatment quality are also unknown.

The literature on treatment quality usually adopts two popular methods. The first method uses patient satisfaction as a measure of treatment quality (Al-Abri and Al-Balushi, 2014; Fenton et al., 2012). However, this measure is subjective and not comparable between different patients. The second method aggregates different objective measures (mortality rate, reinfection rate, average number of days of hospitalisation, medical malpractice rate, etc.) into a single index to reflect treatment quality (Jha et al., 2005; Schreyögg et al., 2011; Nolan, 2001). The problem with this method is to find an appropriate aggregation method.

This chapter applies the environmental performance index (EPI) discussed in Essay 2 to investigate the performance of treatment quality. In this chapter, I call this index the treatment quality performance index (TQI). I analyse the treatment quality performance of Chinese hospitals from different versions of TQI: the output-oriented TQI, the input-oriented TQI and the Malmquist TQI. The performance of treatment quality is a concern of the third objective of the new healthcare reform of China and is rarely studied using the production approach. The final consequences of increasing the number of private hospitals and the amount of government subsidies are generally unknown, as mentioned earlier. Therefore, I estimate the effectiveness of these two measures to evaluate whether their coexistence can be justified. The next section explains the methodology of the treatment quality index.

4.2 Treatment Quality Performance Index

The evaluation of the performance of treatment quality is similar to the evaluation of environmental performance. Both concepts seek to minimise undesirable outputs and maximise desirable outputs. In Essay 2 (Chapter 3), I discuss the environmental performance index used to measure the ability to produce desirable outputs against undesirable outputs as by-products. As deaths during treatment can be treated as undesirable outputs and medical services as desirable outputs, I can evaluate treatment quality based on the environmental performance index. In environmental performance, undesirable outputs are pollutants, while in the performance of treatment quality, undesirable outputs are deaths during treatment. In addition, in environmental performance, desirable outputs are industrial outcomes, while in the performance of treatment quality, desirable outputs are medical services. Improving the performance of treatment quality of hospitals involves reducing the mortality rate of different types of medical services.

I treat each hospital as a production unit, similar to the production definition in Essay 2, with the vector of medical services $g = (g_1, \dots, g_M)$ and the input vector $x = (x_1, \dots, x_N)$. The production of desirable outputs is accompanied by the vector of deaths during treatment $b = (b_1, \dots, b_L)$. The technology is modelled by hospitals, using inputs to provide medical services with deaths as by-products: $\Im = \{(x, g, b): x\ can\ produce\ (g, b)\}$. Similar to Essay 2, there are three types of treatment quality index (TQI): the output-oriented TQI, the input-oriented TQI and the Malmquist TQI. I use these three indices to measure the performance of treatment quality of Chinese hospitals.

The output-oriented TQI of the investigated output vectors $(g^{\mathrm{t}}, b^{\mathrm{t}})$ relative to the output vectors (g^0, b^0) with respect to the reference vector (x^r, g^r, b^r) and the technology in t is

$$TQI^{out}(\, g^{\mathrm{t}}, b^{\mathrm{t}}; g^0, b^0; x^r, g^r, b^r) = \frac{Q_g^{out}(g^t, g^0; x^r, b^r | \Im^t)}{Q_b^{out}(b^{\mathrm{t}}, b^0; x^r, g^r | \Im^t)}. \qquad (64)$$

The input-oriented TQI of the output vectors (g^t, b^t) relative to the output vectors (g^{t-1}, b^{t-1}) with respect to the input reference vector x^r and the technology in t is

$$TQI^{in}(\, g^{\mathrm{t}}, b^{\mathrm{t}}; g^{t-1}, b^{t-1}; x^r) = \frac{Q_g^{in}(g^{\mathrm{t}}, g^{t-1}; x^r | \Im_b^t)}{Q_b^{in}\left(b^{\mathrm{t}}, b^{t-1}; x^r | \Im_g^t\right)}. \qquad (65)$$

The Malmquist TQI with desirable output-oriented distance function and the Malmquist TQI with undesirable output-oriented distance function with respect to the unrestricted input production set $\Im_i^t$ are

$$MTQI_g^t(g^t, b^t; g^{t-1}, b^{t-1}) = \frac{D_g(g^t, b^t | \Im_i^t)}{D_g(g^{t-1}, b^{t-1} | \Im_i^t)}. \qquad (66)$$

and

$$MTQI_b^t(g^t, b^t; g^{t-1}, b^{t-1}) = \frac{D_b(g^t, b^t | \Im_i^t)}{D_b(g^{t-1}, b^{t-1} | \Im_i^t)}. \qquad (67)$$

In general, the treatment quality index indicates the quantity of desirable outputs (treatment quantity index) produced per unit of undesirable outputs (mortality quantity index). The performance of treatment quality improves when the TQI score increases. I evaluate the performance of treatment quality of Chinese hospitals using these three indices and compare the results.

4.3 Treatment Quality Performance in the Chinese Hospital Sector

4.3.1 *Data source*

To evaluate the performance of treatment quality in the Chinese hospital sector, I use a panel dataset of 31 Chinese provinces (exclude Taiwan, Hong Kong, Macau, and Tibet) for the 2009-2014 period collected from the China Public Health Statistical Yearbooks published by the National Bureau of Statistics of China. There are six desirable outputs, five undesirable outputs and six inputs. Desirable outputs are the outcomes of medical services. The six desirable outputs are inpatient visits, outpatient visits, observation visits, babies born in hospitals, physical examinations and surgeries. For inputs, I divide labour into two types: doctors and other medical staff. In addition, the number of beds is used as a proxy for hospital capacity. Therefore, the three inputs are doctors, other medical staff and beds.

In this chapter, undesirable outputs are the five types of deaths during treatment: observation visit deaths, inpatient visit deaths, contagious disease deaths, pregnant women deaths and infant deaths. For deaths during treatment, the literature on the hospital sector usually includes ‘amenable deaths’, which are preventable. The mortality rate is adjusted for different cases of illness. However, my dataset does not contain such information. I follow the approach of Hu et.al. (2012) to adjust undesirable outputs based on provincial life expectancy. Let $life^t$ be the national average life expectancy at time *t,* and $life_k^t$ be the provincial life expectancy at time *t* for province *k*. I define the risk adjustment factor for province *k*: $\delta_k^t = life_k^t/life^t$. A higher value of δ_k indicates a longer life expectancy. In terms of longer life expectancy, more deaths are inevitable and fewer deaths are preventable. Thus, I adjust the mortality due to undesirable outputs b_k^t to the risk adjustment factor δ_k. In the following computation of the treatment quality index, I use the risk-adjusted mortality $\widetilde{b_k^t} = b_k^t/\delta_k^t$ as undesirable outputs. Table 2 lists the definitions and descriptive statistics of outputs and inputs. Note that the undesirable outputs are already adjusted to the risk adjusted mortality rate.

Table 2: Definitions and descriptive statistics of outputs and inputs

Variable	Definition	Mean	*SD*	Min	Max
Desirable outputs					
OPV	Outpatient visits (hundred million)	2.00	1.58	0.09	7.59
OBV	Observation visits (million)	2.12	1.79	0.02	8.47
SUG	Surgeries (million)	11.21	8.46	0.17	47.00
BABY	Babies born (million)	4.75	3.63	0.36	16.34
PHE	Physical examinations (hundred million)	1.07	0.86	0.11	3.93
IPV	Inpatient visits (million)	0.54	0.37	0.14	1.50
Undesirable outputs (Risk-adjusted mortality)					
OBD	Observation visit deaths (risk adjusted)	1129.2	1079.9	0	4776
IPD	Inpatient visit deaths (thousand)	18.09	14.44	0.11	68.08
INFD	Infant deaths (thousand)	2.98	2.02	0.17	8.15
COTGD	Contagious disease deaths (thousand)	0.51	0.62	0.02	3.42
PGD	Pregnant women deaths	76.6	57.2	1.9	226
Inputs					
DOCTOR	Number of doctors (thousand)	83.4	52.3	4	231.8
OTH	Number of other medical staff (thousand)	124.4	78.7	4.5	365.2
BED	Number of beds in hospitals (thousand)	176.7	112.6	8.4	500.6

From Table 2, I cannot obtain information on the change over time of desirable outputs and undesirable outputs. In Table 3, I list the annual average of desirable outputs and undesirable outputs. Thus, I can see the trend over time of different types of medical services and deaths during treatment.

Table 3: Annual average of desirable outputs and undesirable outputs

Variable	2009	2010	2011	2012	2013	2014	$\frac{2014}{2009}$
OPV	1.67	1.78	1.92	2.11	2.25	2.34	140%
OBV	2.18	2.17	2.27	2.10	2.07	1.93	88%
SUG	7.42	9.26	11.09	11.84	12.53	12.03	162%
BABY	4.29	4.56	4.92	5.75	6.17	6.57	153%
PHE	0.85	0.94	1.05	1.19	1.28	1.41	166%
IPV	0.45	0.46	0.47	0.50	0.49	0.49	110%
OBD	1.11	1.10	1.10	1.13	1.19	1.15	104%
IPD	14.53	15.46	16.58	17.88	20.88	23.25	160%
INFD	3.44	3.23	2.96	2.94	2.70	2.63	76%
COTGD	0.48	0.46	0.50	0.55	0.53	0.52	109%
PGD	94.08	89.09	73.38	70.03	67.89	65.17	70%

From Table 3, I find that most desirable outputs increased between 10% and 62% during the investigation period, except for observation visits. In addition, only one undesirable output, inpatient deaths, increased by approximately 60%, while observation deaths and contagious deaths increased only slightly. Conversely, infant deaths and pregnant women deaths decreased by about 30%. Overall, I can see the trend over time, with an increase in desirable outputs and a decrease in undesirable outputs between 2009 and 2014. Thus, the performance of treatment quality improved during the investigation period. In the next section, I use the different versions of TQI to confirm the conclusion regarding the improvement of the performance of treatment quality.

4.3.2 *Empirical frontier and reference vectors*

There are 31 provinces in this dataset. Suppose that hospital *k* at time *t* produces the vector of desirable outputs, $g_k^t = (g_{k1}^t, \dots, g_{kM}^t)$, using the input vector, $x_k^t = (x_{k1}^t, \dots, x_{kN}^t)$. The production of desirable outputs is accompanied by the vector of undesirable outputs, $b_k^t = (b_{k1}^t, \dots, b_{kL}^t)$. Thus, $K = 31$, $M = 6$, $N = 5$ and $L = 3$. I adopt the production frontier introduced by Kuosmanen (2005). This empirical production technology is convex with strongly disposable inputs, strongly disposable desirable outputs and jointly weakly disposable of desirable outputs and undesirable outputs. I have three types of empirical frontiers. The first frontier is the *grand production set* $\mathfrak{I}^*$ constructed from all observed inputs and outputs in the dataset. Following the discussion in Essay 2, the output-oriented TQI can only apply the grand frontier, which contains all observations from the first year to the last year. It is estimated as follows:

$$\mathfrak{I}^* := \left\{ \begin{array}{c} (x; g, b): \sum_{t=1}^{T} \sum_{k=1}^{K} z_k^t g_k^t \geqq g, \\ \sum_{t=1}^{T} \sum_{k=1}^{K} z_k^t b_k^t = b, \\ \sum_{t=1}^{T} \sum_{k=1}^{K} (z_k^t + u_k^t) x_k^t \leqq x, \\ u_k^j \geqq 0, z_k^j \geqq 0, \\ for\ k = 1, \cdots, K; t = 1, \cdots, T \end{array} \right\}. \tag{68}$$

The second frontier is a production set in which the technology at time *t* is constructed from the observed data at time *t* and before. This frontier is called the sequential frontier $\mathfrak{I}^{ts}$ in Essay 2. Under constant returns to scale, the empirical sequential frontier at time $t = 1, \dots, T$ is

$$\mathfrak{I}^{ts} := \left\{ \begin{array}{c} (x^t; g^t, b^t) \colon \sum_{j=1}^{t} \sum_{k=1}^{K} z_k^j \, g_k^j \geqq g^t, \\ \sum_{j=1}^{t} \sum_{k=1}^{K} z_k^j b_k^j = b^t, \\ \sum_{j=1}^{t} \sum_{k=1}^{K} \left(z_k^j + u_k^j\right) x_k^j \leqq x^t, \\ u_k^j \geqq 0, z_k^j \geqq 0, \\ for\ k = 1, \cdots, K; t = 1, \cdots, T \end{array} \right\}. \tag{69}$$

Thus, when the sequential frontier is used, the production frontier in each time is likely to be different from previous frontiers. In contrast, the production frontier is the same over the study period for the grand production set. The following two reference vectors discussed in Chapter 3 are used in the computation.

Min-max reference vectors: $R_{min-max} := \left(x^{max}, g^{min}\right)$. These reference vectors are generally adopted in the empirical applications of TQI in the current literature.

Current-lag reference vectors: For hospital $k = 1, \cdots, K$, data from the previous year $R^{lag} := (x_k^{t-1}, g_k^{t-1}, b_k^{t-1})$ and the current year $R^{cur} := (x_k^t, g_k^t, b_k^t)$ are used to compute two values. The final TQI value is the geometric mean of these two numbers.

The TQI with the current-lag reference vectors shows the relative change in treatment quality from the previous year to the current year in the same hospital. I want to modify the TQI to compare it across different units and different periods and let the TQI start from the first year. I use the following two steps to modify the TQI with the ‘current-lag’ reference vectors.

1) The TQI of the first year with the current-lag reference vectors is replaced by the first-year average reference vectors: $R^{fa} := \left(\frac{\sum_{k=1}^{K} x_k^1}{K}, \frac{\sum_{k=1}^{K} g_k^1}{K}, \frac{\sum_{k=1}^{K} b_k^1}{K}\right)$. In this case, I can compute the TQI score for the first year. I use the current-lag reference vectors from the second year.

2) Starting from the second year, the TQI score with the current-lag reference vectors is multiplied by the TQI score from the previous year to adjust the TQI for comparison between different units, $\text{TQI}_t = \prod_{i=1}^{t-1} TQI_i\,, when\ t > 1.$ In this case, all TQI scores refer to the relative change of the first-year average. As a result, the performance of treatment quality is comparable across the panel data.

In the next section of this chapter, the TQI score with the current-lag reference vectors is adjusted using this method. Next, I compute the TQI scores from different frontiers (grand frontier and sequential frontier), different reference vectors (min-max reference vector and current-lag reference vector) and different approaches (input-oriented approach, output-oriented approach, and Malmquist approach). For the grand frontier and the sequential frontier, I compute the output-oriented TQI and input-oriented TQI. In each orientation, I compute using the min-max and the current-lag reference vectors. As I have two approaches (the input-oriented TQI and the output-oriented TQI), for each approach, I use two frontiers: the grand frontier and the sequential frontier. In addition, I use two reference vectors: the min-max reference vectors and the current-lag reference vectors. Moreover, I compute the environmental performance of the Malmquist TQI. In total, I have $2 \times 2 \times 2 + 2 = 10$ TQI versions for each province in one year. The TQI scores with different reference vectors have different meanings.

The TQI score with the min-max reference vectors shows the relative environmental performance by comparing the ratio of observed desirable outputs to observed undesirable outputs with the ratio of minimum desirable outputs to minimum undesirable outputs. The TQI score with the current-lag reference vectors shows the relative environmental performance by comparing the ratio of observed desirable outputs to observed undesirable outputs with the ratio of desirable outputs from the previous/current year to undesirable outputs from the previous/current year. The input-oriented TQI and the output-oriented TQI are constructed from the quantity index with the reference vectors of desirable outputs and undesirable outputs. Thus, they can use the min-max reference vectors with minimum outputs and maximum inputs as the reference. In contrast, the Malmquist TQI is constructed from the Malmquist productivity index with the pair of desirable outputs and undesirable outputs. Thus, the min-max reference vectors cannot be used for the Malmquist TQI. The list of different EPI versions is shown in Table 4.

Table 4: All versions of TQI

Frontier	Reference	Input-oriented TQI	Output-oriented TQI	Malmquist TQI
Sequential frontier	min-max	V1	V5	NA
	current-lag	V2	V6	V9
Grand frontier	min-max	V3	V7	NA
	current-lag	V4	V8	V10

4.3.3 TQI scores in different versions

Table 3 summarises the average TQI scores for the entire country in 10 TQI versions from 2009 to 2014. For TQI with the current-lag reference vectors (V2, V4, V6, V8 and V9), the TQI scores for the first year are not available. For the output-oriented TQI with the current-lag reference vectors (V6 and V8), there are 161 infeasible cases with the grand frontier and 175 infeasible cases with the sequential frontier out of 184 observations. In addition, the output-oriented TQI with the min-max reference vectors (V5) has 37 infeasible cases out of 184 observations. The average TQI scores are not available from 2009 to 2014. As most cases of V6 and V8 are infeasible, the average scores are also unavailable, while the average score of V5 only reports the feasible cases. The averages TQI scores of the different versions are presented in Table 5,

Table 5: Average TQI scores in different versions

TQI	2009	2010	2011	2012	2013	2014	2009-2014
V1	0.91	1.01	1.07	1.18	1.25	1.27	140%
V2	0.77	0.82	0.84	0.88	0.94	0.99	117%
V3	0.99	1.07	1.14	1.20	1.29	1.27	128%
V4	0.77	0.84	0.90	0.94	1.01	0.99	128%
V5	72.5	60.0	51.1	51.1	50.4	49.5	68%
V6	NA	NA	NA	NA	NA	NA	NA
V7	43.0	44.9	48.3	49.5	53.3	54.7	128%
V8	NA	NA	NA	NA	NA	NA	NA
V9	1.04	1.04	1.14	1.23	1.28	1.33	128%
V10	1.06	1.08	1.18	1.26	1.31	1.33	125%

From Table 5, I obtain the average TQI scores of different versions. The following conclusions can be drawn from the table above. First, V6 and V8 are infeasible in most cases, with both versions using the current-lag reference vectors in the output-oriented TQI. Therefore, I can conclude that I cannot use the current-lag reference vectors in the output-oriented TQI. Second, the average TQI scores in most versions increase during the investigation period (except V5). This shows that overall, the TQI scores of the Chinese hospital sector increased between 2009 and 2014. Third, the results of the sequential frontier and the grand frontier in the output-oriented TQI go in the opposite direction. In V7 with the grand frontier setting, the performance of treatment quality improved by approximately 28% between 2009 and 2014. Conversely, in V5 with the sequential frontier setting, the performance of treatment quality decreased by about 32% between 2009 and 2014. This shows that the output-oriented TQI is not only an infeasibility problem, but it is also sensitive to the reference frontier. Finally, in most indices, the performance of treatment quality improved by about 30%. Thus, I can conclude that the performance of treatment quality improved by about 30% between 2009 and 2014.

The results of the evaluation of the different versions of TQI are robust and consistent with the data. In Table 5, I compare the average dynamic changes of desirable outputs and undesirable outputs $\frac{G_t}{B_t}$ for the first year and the change over time of different TQI versions for the first period $\frac{TQI_t}{TQI_1}$. I normalise all of the indices and first-year average data to 1. The average dynamic changes of desirable outputs and undesirable outputs are computed from the following equation

$$\frac{G_t}{B_t} = \frac{\frac{1}{6}\sum_{i=1}^{6}\frac{g_i^t}{g_i^1}}{\frac{1}{5}\sum_{i=1}^{5}\frac{b_i^t}{b_i^1}}. \tag{70}$$

I compare the difference between the average dynamic changes of desirable outputs and undesirable outputs $\frac{G_t}{B_t}$ and the dynamic changes of TQI, as $\frac{G_t}{B_t}$ can show the change in quality performance. If my measures can also show the performance of treatment quality, the correlation of the two should be strong. The dynamic changes of different TQI versions and the correlation are presented in Table 6.

Table 6: Average TQI scores in different versions

TQI_t/TQI_1	2009	2010	2011	2012	2013	2014	Correlation
V1	1.00	1.11	1.18	1.30	1.37	1.40	0.96
V2	1.00	1.06	1.09	1.14	1.22	1.29	0.89
V3	1.00	1.08	1.15	1.21	1.30	1.28	0.96
V4	1.00	1.09	1.17	1.22	1.31	1.29	0.97
V5	1.00	0.83	0.70	0.70	0.70	0.68	-0.97
V6	NA	NA	NA	NA	NA	NA	NA
V7	1.00	1.04	1.12	1.15	1.24	1.27	0.93
V8	NA	NA	NA	NA	NA	NA	NA
V9	1.00	1.00	1.10	1.18	1.23	1.28	0.92
V10	1.00	1.02	1.11	1.19	1.24	1.25	0.95
G_t/B_t	1.00	1.10	1.23	1.28	1.30	1.32	1.00

V1 to V4 are the input-oriented TQI, V5 to V8 are the output-oriented TQI and V9 and V10 are the Malmquist TQI. From the correlation, the input-oriented TQI and the Malmquist TQI show a strong correlation with $\frac{G_t}{B_t}$. This indicates that the input-oriented TQI and the Malmquist TQI are good measures of the performance of treatment quality. However, the output-oriented TQI presents an infeasibility problem and is sensitive to the reference frontier. The output-oriented TQI with the grand frontier and the min-max reference vectors still shows a strong correlation with $\frac{G_t}{B_t}$. In summary, I suggest that the input-oriented TQI and the Malmquist TQI are good measures of the performance of treatment quality. However, for the moment, I cannot determine which index is the best for measuring the performance of treatment quality. The only difference is that the input-oriented TQI is constructed from the quantity index, while the Malmquist TQI is constructed from the productivity index. In the following analysis, I use the output-oriented TQI, the input-oriented TQI and the Malmquist TQI in a robustness test.

After comparing the average dynamic changes in the performance of treatment quality, the next step is a cross-sectional comparison. This is also an important issue as China is a developing country whose development is unbalanced across different regions. In China, the economic gap between coastal regions and non-coastal regions is significant. Hu et al. (2012) also discuss the difference between coastal regions and non-coastal regions in terms of hospital efficiency. They find that hospitals in coastal regions are significantly more efficient than those in non-coastal regions. Here, I investigate the difference in treatment quality between coastal and non-coastal regions. Table 6 shows the average TQI scores for each year in coastal and non-coastal regions for the output-oriented TQI, the input-oriented TQI and the Malmquist TQI. As the difference between each approach is small, I retain only the results of one version of each TQI in the following analysis. For the output-oriented TQI, I use the grand frontier with the min-max reference vectors. For the input-oriented TQI and the Malmquist TQI, I use the sequential frontier with the 'current-lag' reference vectors. I use the sequential frontier because the grand frontier only contains information on the first-year technology of the sequential frontier. Table 7 presents the average TQI scores of different versions in coastal and non-coastal regions. There are two regional groups: coastal regions and non-coastal regions.

Table 7: Regional average TQI scores from 2009 to 2014

Version	TQI^{out}		TQI^{in}		$MTQI$	
Year	Coast	Non-coast	Coast	Non-coast	Coast	Non-coast
2009	64.94	44.64	0.94	0.94	1.29	1.03
2010	71.36	46.22	1.09	1.06	1.29	1.06

2011	75.30	47.42	1.25	1.21	1.31	1.07
2012	68.74	49.07	1.64	1.45	1.34	1.11
2013	73.73	54.84	2.24	2.04	1.36	1.19
2014	78.38	54.71	3.14	2.72	1.36	1.23
average	72.08	49.48	1.72	1.57	1.32	1.11

All TQI scores for coastal regions are better than non-coastal regions in terms of treatment quality. This shows that the performance of treatment quality of coastal regions is better than that of non-coastal regions. Therefore, the Chinese government should focus on ways to improve treatment quality in non-coastal regions. However, this table cannot show the reasons for the gap between coastal and non-coastal regions. In the next section, I explore the determinants of treatment quality in different provinces.

4.4 Determinants of the Treatment Quality Index

4.4.1 *Data Description and Regression Model*

One of the objectives of this chapter is to examine the effects of market-oriented strategies and government-oriented strategies on treatment quality. The performance of treatment quality is measured by the treatment quality index in Section 4. I use the ratio of private hospitals ($MRKRT$) as an indicator of marketisation, which is a measure of market-oriented policies. A larger value of $MRKRT$ means increased market competition between hospitals. The ratio of hospital revenue from government subsidies ($SBSDYRT$) is used to reflect the importance of government resources in the operation of hospitals, which is a measure of government-oriented policies. A larger value of $SBSDYRT$ indicates more government influence and less competitive pressure on hospital administrators. Following Hu et al. (2012), I include six other control variables in my regression model: $BESTRT$, PER_GDP, $URBANRT$, $POPDENSE$, $MEDCOSM$ and $COAST$.

$BESTRT$ is the ratio of first-class hospitals to all hospitals in each province. In China, each hospital is classified at a different level based on its size and the quality of its treatment. First-class hospitals are called ‘San-jia’ hospitals. According to The Measures for the Administration of Hospital Grades published by the Ministry of Health in China, to be a San-jia hospital, a hospital must meet the following requirements: 1) have more than 500 beds for inpatient services; 2) have more than 25 departments for medical services and research in the given list; 3) have more than 1.04 medical technical staff and 0.4 nurse for each bed; 4) have more than two associate professors in each department; 5) have more than 60 types of medical equipment in the given list; and 6) meet other fixed asset and space size requirements. Until December 2017, only 1,331 hospitals in the country

Were classified as San-jia hospitals. These San-jia hospitals had the best equipment, the most advanced medical technologies and the most efficient medical services in China. I control the ratio of first-class hospitals as it refers to the allocation of medical resources in each province, which can affect treatment quality. PER_GDP is the real GDP per capita based on its value in 2008. Higher real GDP per capita refers to a higher level of income. I expect it to affect treatment quality as a Well-developed economy can help the development of the hospital sector. $URBANRT$ is the ratio of urban residents in each province. Medical resources are usually centralised in cities. Therefore, in the process of urbanisation, more people benefit from the high quality of treatment of hospitals in large cities. $POPDENSE$ is the population density. Population density has two effects on treatment quality. First, a higher population density can increase the demand for medical services. Thus, medical services are more needed in densely populated regions, which may help develop high-quality services. Second, a higher population density can increase the pressure of the health administration. Indeed, contagious diseases are easier to transmit and the probability of an accident may also increase. $MEDCOSM$ is the ratio of medical expenditure to individual consumption. I control this variable as a proxy for the price of medical services. Usually, a higher quality of services may require a higher price. $COAST$ is the regional dummy to indicate if this province is a coastal region. In the previous section, I find that coastal regions have better treatment quality than non-coastal regions. By controlling these environmental variables, I examine whether the regional difference still persists.

Although I adopt the same control variables as Hu et al. (2012), the relationships between the dependent variable and the control variables are different. In particular, the dependent variable in this chapter is treatment quality, whereas Hu et al. (2012) use technical efficiency. The variables and their definitions are presented in Table 8. The data for the regressors are collected from the China Public Health Statistical Yearbooks between 2009 and 2014.

Table 8: Definitions and descriptive statistics of the regression variables

Variable	Definition	Mean(*SD*)	min	max
MRKRT	Market rate: ratio of the number of private hospitals (%)	34.16(14.55)	3	69.42
SBSDYRT	Subsidy rate: ratio of revenue in the health sector from government subsidies (%)	16.71(7.47)	8.23	53.79
BESTRT	Ratio of first-class hospitals to all hospitals (%)	4.41(2.24)	0.97	10.98
POPDENSE	People per square kilometre	371(418)	2	2303
PER_GDP	GDP per capita based on 2008 (thousand)	37.35(19.86)	10.26	99.6
URBANRT	Ratio of urban residents in the population	37.94(16.79)	16.14	90.03
MEDCOSM	Ratio of medical expenditure to consumption (%)	7.37(1.81)	2.53	11.86
COAST	Regional dummy variable: coastal areas			

The regression model is as follows:

$$\begin{aligned}\mathrm{TQI}_{i,t} = \beta_0 + \beta_1 MRKRT_{i,t} + \ & \beta_2 SBSDYRT_{i,t} + \ \beta_3 BESTRT_{i,t} \\ & + \beta_4 \ln(PER_{GDP_{i,t}}) \ + \ \beta_5 URBANRT_{i,t} + \ \beta_6 \ln(POPDENSE_{i,t}) \\ & + \beta_7 MEDCOSM_{i,t} + \beta_8 COAST_{it} + \varepsilon_{i,t}. \qquad (71)\end{aligned}$$

The first two explanatory variables are the focus of this chapter. If β_1 is positive, then a larger value of $MRKRT$ will increase the value of TQI. As $MRKRT$ is a proxy for market-oriented policies, this means that marketisation helps promote treatment quality in the hospital sector. Conversely, if β_2 is positive, then a larger value of $SBSDYRT$ will increase the value of TQI. As $SBSDYRT$ is a proxy for government-oriented policies, this means that government intervention can improve the treatment quality of the hospital sector.

Their coefficients are expected to be positive. For example, the ratio of first-class hospitals ($BESTRT$) is an indicator to show the quality of the hospital sector in each province. A higher value of $BESTRT$ indicates that a province has superior medical equipment and knowledge. I expect $\beta_3 > 0$ because first-class hospitals should provide better medical services and cause fewer deaths. Furthermore, real GDP per capita (PER_GDP) indicates the income level of each province. People with higher income should be healthier. Ceteris paribus, the mortality rate of high-income patients should be lower. Thus, I expect $\beta_4 > 0$. In the next section, I present the empirical results of different versions of TQI and different estimation models.

4.4.2 Estimation results

For panel data regression, I use the period fixed effects model. The estimated results of the Regression Equation (71) are reported in Model (1) of Table 9. Another consideration is the potential problem of endogeneity between the performance of treatment quality (represented by TQI) and market forces and government subsidies (represented by $MRKRT$ and $SBSDY$). There is also a potential problem of reverse causality, in which treatment quality may affect market competition or government subsidies because the government may want to subsidise more hospitals with advanced medical technologies, leading to higher treatment quality. To solve the endogeneity problem, I use the two-stage least squares (2SLS) regression and the instrumental variables (IV) regression. The instrumental variables are the lag terms of *MRKRT* and *SUBSDRT*. The lag terms of *MRKRT* and *SUBSDRT* meet the requirements of the instrumental variables. First, they are correlated, as there is autocorrelation in time series data. Second, these lag terms are not theoretically correlated with the current error term in the Regression Equation (71). I check the effectiveness of my IV regression. The values of the first-stage F-test are 37 and 45. Both support the validity of the lag terms of $MRKRT$ and $SUBSDRT$ as IV. The estimated results of the Regression Equation (14) with instrumental variables are reported in Model (3) of Table 9.

Table 9: Determinants of TQI from different models

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
	$\mathrm{TQI}_{out}\ OLS$	$\mathrm{TQI}_{out}\ IV$	$\mathrm{TQI}_{in}\ OLS$	$\mathrm{TQI}_{in}\ IV$	$\mathrm{TQI}_{mq}\ OLS$	$\mathrm{TQI}_{mq}\ IV$
Constant	-211.16***	-329.77***	1.33	0.91	1.83	1.08
	(53.5)	(66.42)	(1.16)	(1.49)	(2.87)	(3.72)
MRKRT	**0.25****	**0.23***	**0.01*****	**0.02*****	**0.02*****	**0.016*****
	(0.11)	(0.14)	(0.003)	(0.003)	(0.006)	(0.005)
SUBDY	**1.73*****	**2.86*****	**0.02*****	**0.04*****	**0.02*****	**0.036*****
	(0.33)	(0.5)	(0.007)	(0.01)	(0.0017)	(0.003)
BESTRT	-1.76***	2.06**	-0.007	0.02	-0.013	-0.009
	(0.72)	(0.86)	(0.015)	(0.02)	(0.039)	(0.05)
PER_GDP	17.57***	23.06***	0.35***	0.43***	0.25***	0.36***
	(0.72)	(6.47)	(0.12)	(0.14)	(0.09)	(0.06)
MEDCOM	0.9	1.45*	0.04***	0.07***	0.04***	0.05***
	(0.8)	(0.95)	(0.017)	(0.02)	(0.01)	(0.01)
URBANRT	-0.35***	-0.35***	0.01***	0.01***	0.04***	0.04***
	(0.16)	(0.19)	(0.003)	(0.003)	(0.007)	(0.007)
POPDEN	8.71***	13.82***	-0.03	-0.001	0.12	0.05
	(1.78)	(2.43)	(0.04)	(0.05)	(0.09)	(0.13)
COAST	3.19	2.28	0.34***	0.45***	0.33***	0.50***
	(3.57)	(4.13)	(0.07)	(0.09)	(0.08)	(0.12)
N	186	155	186	155	186	155
$\widetilde{R}^2$	0.34	0.43	0.32	0.35	0.64	0.71

Notes: Numbers in parentheses are standard errors.

*, ** and *** indicate significance at the 10%, 5% and 1% levels.

From Model (1) to Model (6), all estimation results of $MRKRT$ and $SBSDY$ consistently show that marketisation and government subsidies help significantly improve the performance of treatment quality. The results of the IV regression and the simple OLS regression for the models show no significant difference, indicating that my models do not have a serious endogeneity problem. For the goodness of fit, the adjusted R-squared in the Malmquist TQI model is about twice that of the input-oriented TQI model and the output-oriented TQI model. As a result, the Malmquist TQI is better than the output-oriented TQI and the input-oriented TQI in this dataset. For the control variables, different models generate different results, although only PER_GDP is consistently significantly positive. This shows that a higher level of economic development can improve the performance of treatment quality in the hospital sector. Compared with Model (1) and Model (2), Model (3) to Model (6) show consistent results for other control variables. This indicates that the input-oriented TQI and the Malmquist TQI are more robust than the output-oriented TQI. In the following analysis, I use the estimation results of the input-oriented TQI and the Malmquist TQI. From Model (3) to Model (6), PER_GDP, $URBANRT$, $MEDCOSM$ and $COAST$ are significantly positive. This reveals that real GDP per capita and urbanisation can improve the performance of treatment quality. In addition, higher quality medical services require a higher price, which explains the significant positive relationship between medical consumption and TQI. Moreover, I find that coastal regions still perform significantly better in terms of treatment quality than non-coastal regions after excluding the effects of other environmental factors. $BESTRT$ and

$URBANRT$ have no significant effect in Model (3) to Model (6). These results are unexpected, as they show that more first-class hospitals cannot improve the performance of treatment quality.

4.4.3 *Analysis of the Regression Results*

The estimated results of the six models show that marketisation and government subsidies can improve treatment quality in the hospital sector in China. The estimated coefficients of $MRKRT$ and $SBSDY$ in Table 9 are significant at 1%. Specifically, when the ratio of private hospitals ($MRKRT$) increases by 1 percentage point, TQI increases by about 0.02. When the ratio of revenue from government subsidies ($SBSDY$) increases by 1 percentage point, TQI increases by about 0.04. From Table 9, the means of $MRKRT$ and $SBSDYRT$ are 34.16% and 16.71%, respectively. Both marketisation and the level of subsidies are at a low level. Therefore, there is great potential to improve treatment quality by increasing the number of for-profit hospitals and increasing subsidies to the hospital sector.

For the control variables, as the output-oriented TQI has different results from the input-oriented TQI and the Malmquist TQI, I use the results of Model (3) to Model (5) to explain the control variables, because the results of the control variables are not consistent from Model (1) to Model (6). The estimation results and the explanations are only for reference. The results of the control variables are not robust and strong enough for policy implications. As expected in the previous section, the estimated coefficients of real GDP per capita ($ln(PER\_GDP)$) are significantly positive. This result confirms that higher GDP per capita can increase the value of TQI. Thus, economic growth should improve treatment quality in the healthcare sector. The estimated coefficients of the ratio of urban residents in the population ($URBANRT$) are also significantly positive. This means that a higher rate of urbanisation has a positive effect on treatment quality. Medical resources are usually centralised in cities. Indeed, in the process of urbanisation, more people benefit from the high treatment quality of hospitals in large cities. The ratio of medical expenditure to consumption ($MEDCOM$) also has a significant positive effect on the performance of treatment quality. This may be due to the higher cost of providing high-quality medical services. When the performance of treatment quality improves, the hospital sector will charge higher prices to cover additional costs. The regional dummy variable for coastal regions ($COAST$) indicates that after controlling for other variables, there is still a difference between coastal and non-coastal provinces. Thus, the Chinese government should focus more on the balanced development between coastal regions and non-coastal regions with regard to the performance of treatment quality.

4.5 Mixed Policies of Healthcare Reforms in China

The empirical analysis gives several important results. First, the Chinese hospital sector performance improved during the 2009-2014 period in terms of patient treatment. Although the results show a clear upward trend in desirable outputs, undesirable outputs do not follow this trend and remain constant. This indicates a decrease in the mortality rate during treatment. Thus, treatment quality has improved over time. Second, there is a big gap between the coastal regions and non-coastal regions of different provinces. From an equity perspective, the government should try to eliminate this big gap and strike a balance in the development of the hospital sector. Finally, if the ratio of private hospitals ($MRKRT$) and the ratio of revenue from government subsidies ($SBSDY$) can capture the main effects of market-oriented policies and government-oriented policies on the hospital sector, then the coexistence of the two policies should improve treatment quality in the hospital sector in China.

In previous studies, market-oriented strategies and government-oriented strategies are distinct alternatives. Before 2009, healthcare reforms in China Were characterised by market-oriented strategies or government-oriented strategies, but not both. However, the new healthcare reform launched in 2009 in China adopts a mixed policy, strengthening market forces and government intervention at the same time. According to the official documents of the National Health and Family Planning Commission on the healthcare reform published in 2015 and 2016, the Chinese government plans to simultaneously increase the ratio of private hospitals to market share and financial support to the hospital sector.

As mentioned in Section 4.1, one of the main goals of the current healthcare reform is to improve treatment quality in the hospital sector. In 2014, the ratio of private hospitals in China was 49% and the ratio of revenue from government subsidies was only 12%. There is room to increase the market share of private hospitals and government subsidies. Hence, the mixed strategy is feasible. My empirical results in Section 4.4 confirm the positive effects of increasing the number of private hospitals and increasing government subsidies on treatment quality. Therefore, improving treatment quality is both feasible and achievable by adopting a mixed strategy.

According to the official report of the Chinese government, one of the main objectives of the government is to promote treatment quality in general. The mixed strategy can help increase treatment quality. First, the market-oriented ownership reform can help improve treatment quality. Indeed, Hadley et.al. (1996) show that financial pressure improves the efficiency of hospitals. Bloom, Propper and Van (2015) also find that competition promotes treatment quality. Therefore, the government should encourage more private investors to invest in the hospital sector to increase market competition. Second, government subsidies should be increased. Duggan (2000) argues that increasing government spending cannot increase healthcare outcomes. However, I find that increasing government spending can improve treatment quality. There are two ways to support the healthcare sector. One way is to support the supply side by investing directly in the creation of public hospitals or directly subsidising hospitals. The other way is to support the demand side by paying for medical insurance or directly subsidising patients. Due to data availability, this study focuses solely on supply-side subsidies to hospitals, including public and private hospitals.

My study recommends the mixed strategy to promote treatment quality as follows: (i) the government should encourage more market competition in the current status of the hospital sector. Strong market competition in the hospital sector can improve treatment quality. (ii) The government has a crucial role to play in providing more support to promote treatment quality. The government should increase its financial support to private hospitals, as public hospitals receive most of the financial support at the current stage.

4.6 Conclusions

This chapter analyses the performance of treatment quality of the hospital sector in China using the treatment quality index (TQI) from 2009 to 2014. Adopting the data envelopment analysis approach, I compute the treatment quality index. I find that the treatment quality index of the hospital sector in China increases during the studied period. The current healthcare reform launched in 2009 has improved the treatment quality of the hospital sector. I also identify different patterns of treatment quality between coastal and non-coastal provinces.

This chapter also reveals the determinants of treatment quality. The current healthcare reform adopts a mixed strategy. If the ratio of private hospitals ($MRKRT$) and the ratio of revenue from government subsidies ($SBSYD$) can capture the main effects of market-oriented policies and government-oriented policies, then my empirical results confirm that the mixed strategy of the current healthcare reform is moving in the right direction to improve treatment quality. The estimated regression equations also indicate that after controlling for other variables, the regional difference between coastal and non-coastal provinces still persists. The government should pay more attention to the balanced development of the hospital sector, particularly to improving treatment quality in non-coastal regions.

The results of this study are subject to the following caveats. First, this chapter only estimates the effects of subsidies without considering other major intervention measures. For an in-depth study of government intervention, a larger dataset containing more information should be used. Second, my treatment quality index does not capture all aspects of the overall quality of a hospital's services. If data are available, more dimensions other than deaths should be included. Third, provincial data are used in this study. However, a dataset with individual hospitals may be more appropriate as hospitals are real production units.

Chapter 5

Summary

This thesis investigated the quality of medical services in the hospital sector from measures and application. Chapter 2 discussed the definition of treatment quality and critical reviewed the measurement of quality assessment. It shows that the environmental performance index can applied to hospital sector as quality assessment tool. Chapter 3 proposed two new environmental performance indices, which can also be used to evaluate the performance of treatment quality in the hospital sector. Chapter 4 used the new measures of the performance of treatment quality to evaluate the quality of the Chinese hospital sector and further investigated the effects of marketisation and government subsidies on treatment quality.

Chapter 2 (Essay 1) defines the quality from favourable and unfavourable dimensions. This chapter classified the quality assessment from production approach and non-production approach. It shows that the environmental performance index can apply to hospital sector. The production approach with weak disposability assumption of undesirable outputs is suitable for assess treatment quality.

Chapter 3 (Essay 2) explicitly highlights the weakness of the output-oriented environmental performance index, which is ill-defined without restrictive conditions (min-max reference vector and the grand frontier). This essay proposes two alternative approaches input-oriented EPI and Malmquist EPI to construct measure environmental performance: replacing output distance functions by input distance functions in the formulas of output quantity indices of the output-oriented EPI, and replacing all output quantity indices by the Malmquist productivity index. It has been shown that both new formulas improve upon the output-oriented EPI in ensuring finite values of EPI when various production frontiers and reference vectors are adopted. Further, based on the Malmquist index, I decompose the change of environmental performance. The decomposition of Malmquist EPI can identify the source of environmental performance change as shifting of technology or the efficiency improvement.

Chapter 4 (Essay 3) analyse the performance of treatment quality in the hospital sector in China using the treatment quality index (TQI) from 2009 to 2014. Adopting the data envelopment analysis approach, I compute the treatment quality index. I show that the treatment quality index of the hospital sector in China increased during the study period. The current healthcare reform launched in 2009 has improved the treatment quality of the hospital sector. I also find different patterns of treatment quality between coastal and non-coastal provinces. This chapter also investigates the determinants of treatment quality. The current healthcare reform adopts a mixed strategy. If the ratio of private hospitals and the ratio of revenue from government subsidies policies, then my empirical results confirm that the mixed strategy of the current healthcare reform is moving in the right direction to improve treatment quality.

References

Adler, N., and J. Berechman, 2001. Measuring Airport Quality from the Airlines' Viewpoint: An Application of Data Envelopment Analysis, *Transport Policy, 8*(3), pp.171-181.

Ahmad, I., Nawaz, A., Khan, S., Khan, H., Rashid, M.A. and Khan, M.H., 2011. Predictors of patient satisfaction. *Gomal Journal of Medical Sciences, 9*(2).

Aiken, L.H., Clarke, S.P., Sloane, D.M. and International Hospital Outcomes Research Consortium, 2002. Hospital staffing, organization, and quality of care: cross-national findings. *International Journal for quality in Health care,* 14(1), pp.5-14.

Akin, J. S., Dow, W. H., Lance, P. M., and Loh, C. P. A., 2005. Changes in access to health care in China, 1989–1997. *Health Policy and Planning*, *20*(2), pp. 80-89.

Aiken, L.H., Sermeus, W., Van den Heede, K., Sloane, D.M., Busse, R., McKee, M., Bruyneel, L., Rafferty, A.M., Griffiths, P., Moreno-Casbas, M.T. and Tishelman, C., 2012. Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. *Bmj, 344,* p.e1717.

Al-Abri, R., and Al-Balushi, A., 2014. Patient satisfaction survey as a tool towards quality improvement. *Oman Medical Journal*, *29*(1), p. 3.

Arjomandi, A., K.H. Dakpo, and J.H. Seufert, 2018. Have Asian Airlines Caught Up With European Airlines? A By-Production Efficiency Analysis, *Transportation Research*, 116A, pp.389-403

Armstrong, M., and Chen, Y., 2009. Inattentive consumers and product quality. *Journal of the European Economic Association*, *7*(2-3), pp. 411-422.

Bai, C. E., Lu, J., and Tao, Z., 2006. The multitask theory of state enterprise reform: Empirical evidence from China. *American Economic Review*, *96*(2), pp. 353-357.

Bjertnaes, O.A., Sjetne, I.S. and Iversen, H.H., 2012. Overall patient satisfaction with hospitals: effects of patient-reported experiences and fulfilment of expectations. *BMJ Qual Saf*, *21*(1), pp.39-46.

Blomqvist, Å., 2009. Health system reform in China: What role for private insurance? *China Economic Review*, *20*(4), pp. 605-612.

Bloom, N., Propper, C., Seiler, S., and van Reenen, J., 2015. The impact of competition on management quality: Evidence from public hospitals. *Review of Economic Studies, 82*(2), pp. 457-489.

Blumenthal, D. and Hsiao, W., 2015. Lessons from the East—China's rapidly evolving health care system. *New England Journal of Medicine, 372*(14), pp.1281-1285.

Brekke, K.R., Siciliani, L. and Straume, O.R., 2011. Hospital competition and quality with regulated prices. *Scandinavian Journal of Economics*, *113*(2), pp.444 - 469.

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

Chen, Z., Liu, X., and Wang, H., 2008. Banish the price regulation of medical services: The next step of medical system reform in China. *Management World, 3*(8), pp. 52-58 (In Chinese).

Cheng, S.H., Yang, M.C. and Chiang, T.L., 2003. Patient satisfaction with and recommendation of a hospital: effects of interpersonal and technical aspects of hospital care. *International Journal for Quality in Health Care, 15*(4), pp.345-355.

Chioveanu, I., 2012. Price and quality competition. *Journal of Economics*, *107*(1), pp. 23-44.

Chung, Y. H., Färe, R., and Grosskopf, S., 1997. Productivity and Undesirable Outputs: A Directional Distance Function Approach. *Journal of Environmental Management, 51*(3), pp. 229-240.

Cleary, P.D. and McNeil, B.J., 1988. Patient satisfaction as an indicator of quality care. *Inquiry*, pp.25-36.

Clever, S.L., Jin, L., Levinson, W. and Meltzer, D.O., 2008. Does doctor–patient communication affect patient satisfaction with hospital care? Results of an analysis with a novel instrumental variable. *Health services research, 43*(5p1), pp.1505-1519.

Coelli, T., Lauwers, L., and Van Huylenbroeck, G., 2005. Formulation of technical, economic and environmental efficiency measures that are consistent with the materials balance condition. In CEPA Working Papers Series (WP062005) (pp. 1–38). School of Economics, University of Queensland: Australia

Considine, T. J., and Larson, D. F., 2006. The environment as a factor of production. *Journal of Environmental Economics and Management, 52*, pp.645–662.

Cui, Q. and Y. Li., 2017. Airline Efficiency Measures under CNG2020 Strategy:

An Application of a Dynamic By-production Model, *Transportation Research, 106*A, pp.130-143.

Dakpo, K.H., P. Jeanneaux, and L. Latruffe., 2016. Modelling Pollution-Generating Technologies in Performance Benchmarking: Recent Developments, Limits and Future Prospects in the Nonparametric Framework, *European Journal of Operational Research, 250*(2), pp.347-359.

Deily, M.E., and McKay N.L., 2006. Cost inefficiency and mortality rates in Florida hospitals, *Health Economics, 15*(4), pp.419–431.

Donabedian, A., 2005. Evaluating the quality of medical care. *The Milbank Quarterly*, *83*(4), pp. 691-729.

Doran, J.W., and Jones, A.J., 1996. Methods for assessing soil quality. Soil Science Society of America Special Publication, vol. 49. *Soil Science Society of America*, Madison, WI. pp. 410

Doran, J.W., Coleman, D.C., Bezdicek, D.F., Stewart, B.A., 1994. Defining Soil Quality for a Sustainable Environment. Soil Science Society of America Special Publication No. 35. *Soil Science Society of America and American Society of Agronomy*, Madison, WI.

Duggan, M. G., 2000. Hospital ownership and public medical spending. *Quarterly*

Journal of Economics, *115*(4), pp. 1343-1373.

Färe, R. and Grosskopf, S., 2003. Nonparametric productivity analysis with undesirable outputs: comment. *American Journal of Agricultural Economics, 85*(4), pp.1070-1074.

Färe, R., and Grosskopf, S., 2009. A comment on weak disposability in nonparametric production analysis. *American Journal of Agricultural Economics, 91*, pp.535–538.

Färe, R., Grosskopf, S., and Hernandez-Sancho, F., 2004. Environmental performance: An index number approach. *Resource and Energy Economics*, *26*(4), pp. 343-352.

Färe, R., Grosskopf, S., and Pasurka, C., 1986. Effects on relative efficiency in electric power generation due to environmental controls. *Resources and Energy, 8*, pp.167–184.

Färe, R., Grosskopf, S., Lovell, C. A. K., and Pasurka, C., 1989. Multilateral productivity comparisons when some outputs are undesirable: A nonparametric approach. *Review of Economics and Statistics, 71*, pp. 90-98.

Färe, R., Grosskopf, S., Lundgren, T., Marklund, P.O., and Zhou, W., 2012. Productivity: Should we include bads? *CERE working papers*.

Färe, R., Grosskopf, S., and Margaritis, D., 2011. Malmquist productivity indexes and DEA. In W. Cooper, L. Seiford, and J. Zhu (Eds.), *Handbook on data envelopment analysis* (pp. 127-149). Boston, MA: Springer.

Färe, R., S. Grosskopf, and P. Roos., 1995. Productivity and Quality Changes in Swedish Pharmacies, *International Journal of Production Economics, 39*(1), pp.137-147.

Feng, J., and Yu, Y., 2008. Medical system reform: The market-oriented policy and government-oriented policy. *World Economics Chapters*, *1*(1), pp. 1-13.

Ferrari A. 2006. Market oriented reforms of health services: a non-parametric analysis. *The Service Industry Journal, 26*(1), pp.1–13.

Field, B., 1994. Environmental economics, an introduction. McGraw-Hwell International Editions, New York. pp. 482

Fenton, J. J., Jerant, A. F., Bertakis, K. D., and Franks, P., 2012. The cost of satisfaction: A national study of patient satisfaction, health care utilization, expenditures, and mortality. *Archives of Internal Medicine*, *172*(5), pp. 405-411.

Forder, J. and Allan, S., 2014. The impact of competition on quality and prices in the English care homes market. *Journal of Health Economics*, 34, pp.73-83.

Førsund, F.R., 2009. Good Modelling of Bad Outputs: Pollution and Multiple-Output Production, *International Review of Environmental and Resource Economics, 3*(1), pp.1-38.

Fu, S., Viard, B. and Zhang, P., 2017. Air quality and manufacturing firm productivity: Comprehensive evidence from China. *MPRA*

Gerald, B., and Gu, X., 1997. Health sector reform: Lessons from China. *Social Science and Medicine*, *45*(3), pp. 351-360.

Gravelle, H. and Sivey, P., 2010. Imperfect information in a quality-competitive hospital market. *Journal of Health Economics*, 29(4), pp.524-535.

Grossman, M., 1972. On the concept of health capital and the demand for health. *Journal of Political economy,* 80(2), pp.223-255.

Gu, X., 2005. China’s healthcare system reform: Status quo and challenges. Social Blue Book – 2006: Analysis and Forecast of Social Situation in China. *Beijing: Social Science Academic Press* (In Chinese).

Hadley, J., Zuckerman, S., and Iezzoni, L. I., 1996. Financial pressure and competition: Changes in hospital efficiency and cost-shifting behavior. *Medical Care*, *34*(3), pp. 205-219.

Hailu, A., and R.G. Chambers., 2012. A Luenberger Soil-Quality Indicator, *Journal of Productivity Analysis, 38*(2), pp.145-154.

Hailu, A. and Veeman, T.S., 2001. Non-parametric productivity analysis with undesirable outputs: an application to the Canadian pulp and paper industry. *American Journal of Agricultural Economics, 83*(3), pp.605-616.

Harrison, M. I., and Calltorp, J., 2000. The reorientation of market-oriented reforms in Swedish health-care. *Health Policy*, *50*(3), pp. 219-240.

He, Y., and Tian, Z., 2008. Government-oriented corporate public relation strategies in transitional China. *Management and Organization Review*, *4*(3), pp. 367-391.

Helderman, J. K., Schut, F. T., van der Grinten, T. E., and van de Ven, W. P., 2005. Market-oriented health care reforms and policy learning in the Netherlands. *Journal of Health Politics, Policy and Law*, *30*(1-2), pp. 189-210.

Herr, A., 2008. Cost and technical efficiency of German hospitals: Does ownership matter? *Health Economics*, *17*(9), pp. 1057-1071.

Heshmati, A., 2002. Quality Adjusted Measures of Services in Public Schools, *European Journal of Operational Research, 136*(3), pp.655-670.

Hirth, R.A., 1997. Competition between for-profit and non-profit health care providers: Can it help achieve social goals? *Medical Care Research and*

Review, 54(4), pp.414-438.

Hollingsworth, B., 2003. Non-parametric and parametric applications measuring efficiency in health care. *Health care management science, 6*(4), pp.203-218.

Hollingsworth, B., 2008. The measurement of efficiency and productivity of health care delivery. *Health Economics*, *17*(10), pp. 1107-1128.

Ho, C.S., 2010. Health reform and de facto federalism in China. *China: An International Journal*, 8(01), pp.33-62.

Hu, H. H., Qi, Q., and Yang, C. H., 2012. Analysis of hospital technical efficiency in China: Effect of health insurance reform. *China Economic Review*, *23*(4), pp. 865-877.

J. Sørensen, R. and Grytten, J., 1999. Competition and supplier‐induced demand in a health care system with fixed fees. *Health Economics, 8*(6), pp.497-508.

Jha, A. K., Li, Z., Orav, E. J., and Epstein, A. M., 2005. Care in US hospitals – The Hospital Quality Alliance program. *New England Journal of Medicine*, *353*(3), pp. 265-274.

Jun, M. and Palacios, S., 2016. Examining the key dimensions of mobile banking service quality: an exploratory study. *International Journal of Bank Marketing, 34*(3), pp.307-326.

Kiadaliri, A.A., Jafari, M. and Gerdtham, U.G., 2013. Frontier-based techniques in measuring hospital efficiency in Iran: a systematic review and meta-regression analysis. *BMC health services research, 13*(1), p.312.

Karlen, D.L., Mausbach, M.J., Doran, J.W., Cline, R.G., Harris, R.F. and Schuman, G.E., 1997. Soil quality: a concept, definition, and framework for evaluation (a guest editorial). *Soil Science Society of America Journal, 61*(1), pp.4-10.

Karlen, D.L., Ditzler, C.A., and Andrews, S.S., 2003. Soil quality: why and how? *Geoderma 114*, pp.145–156.

Kerstens, K., and Van de Woestyne, I., 2014. Comparing Malmquist and Hicks-Moorsteen productivity indices: Exploring the impact of unbalanced vs. balanced panel data. *European Journal of Operational Research*, *233*(3), pp. 749-758.

Kiadaliri, A.A., M. Jafari, and U.-G., Gerdtham., 2013. Frontier-based Techniques in Measuring Hospital Efficiency in Iran: A Systematic Review and Meta-Regression Analysis, *BMC Health Services Research, 13*(1), pp.312.

Kohl, S.G. 1955. Perinatal Mortality in New York City: Responsible Factors. *Cambridge: Harvard University Press.*

Kolstad, J. T., and Kowalski, A. E., 2012. The impact of health care reform on hospital and preventive care: Evidence from Massachusetts. *Journal of Public Economics*, *96*(11), pp. 909-929.

Kuosmanen, T., 2005. Weak disposability in nonparametric production analysis with undesirable outputs. *American Journal of Agricultural Economics*, *87*(4), pp. 1077-1082.

Lauwers, L., and Van Huylenbroeck, G., 2003. Materials balance based modelling of environmental efficiency. In 25th international conference of agricultural economist (pp. 16–22). South Africa.

Lipworth, L., J.A.H. Lee, and J.N. Morris. April–June, 1963. Case Fatality in Teaching and Nonteaching Hospitals, 1956–1959. *Medical Care* 1, pp.71–76.

Liu, J., 2005. Excessive marketization and decentralization – Dual misunderstanding of China's health care reform. *Health Economics Research*, *12*, pp. 3-10 (In Chinese).

Liu, Y., 2004. China's public health-care system: Facing the challenges. *Bulletin of the World Health Organization*, *82*(7), pp. 532-538.

López-Casasnovas, G., and Puig-Junoy, J., 2000. Review of the literature on reference pricing. *Health Policy*, *54*(2), pp. 87-123.

Mahlberg, B., and Sahoo, B. K., 2011. Radial and non-radial decompositions of Luenberger productivity indicator with an illustrative application. *International Journal of Production Economics, 131*, 721–726.

Mullen, K. J., Frank, R. G., and Rosenthal, M. B., 2010. Can you get what you pay for? Pay-for-performance and the quality of healthcare providers. *The Rand Journal of Economics*, *41*(1), pp. 64-91.

Munksgaard, J., Christoffersen, L. B., Keiding, H., Pedersen, O. G., and Jensen, T. S., 2007. An environmental performance index for products reflecting damage costs. *Ecological Economics*, *64*(1), pp. 119-130.

Murty, S., R.R., Russell, and S.B., Levkoff., 2012. On Modelling Pollution-Generating Technologies, *Journal of Environmental Economics and Management, 64*(1), pp.117-135.

Nolan, T., Angos, P., Cunha, A. J., Muhe, L., Qazi, S., Simoes, E. A., and Pierce, N. F., 2001. Quality of hospital care for seriously well children in less-developed countries. *The Lancet*, *357*(9250), pp. 106-110.

Oh, D. H., and Heshmati, A., 2010. A sequential Malmquist-Luenberger productivity index: Environmentally sensitive productivity growth considering the progressive nature of technology. *Energy Economics*, *32*(6), pp. 1345-1355.

Olesen, O.B., and N.C., Petersen., 1995. Incorporating Quality into Data Envelopment Analysis: A Stochastic Dominance Approach, International *Journal of Production Economics, 39*(1-2), pp.117-135.

O'Newell, L., Rauner, M., Heidenberger, K., and Kraus, M., 2008. A cross-national comparison and taxonomy of DEA-based hospital efficiency studies. *Socio-Economic Planning Sciences*, *42*(3), pp. 158-189.

Ozcan, Y. A., Luke, R. D., and Haksever, C., 1992. Ownership and organizational performance: A comparison of technical efficiency across hospital types. *Medical Care*, *30*(9), pp. 781-794.

Pastor, J. T., and Lovell, C. K., 2005. A global Malmquist productivity index. *Economics Letters*, *88*(2), pp. 266-271.

Paul, C. J. M., Ball, V. E., Felthoven, R. G., Grube, A., and Nehring, R. F., 2002. Effective costs and chemical use in United States agricultural production: Using the environment as a "free" input. *American Journal of Agricultural Economics, 84*, pp.902–915.

Pelone, F., Kringos, D.S., Romaniello, A., Archibugi, M., Salsiri, C. and Ricciardi, W., 2015. Primary care efficiency measurement using data envelopment analysis: a systematic review. *Journal of medical systems, 39*(1), p.156.

Peters, E., Dieckmann, N., Dixon, A., Hibbard, J. H., and Mertz, C. K., 2007. Less is more in presenting quality information to consumers. *Medical Care Research and Review*, *64*(2), pp. 169-190.

Pitches, D.W., Mohammed, M.A., and Lilford, R.J., 2007. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC health services research*, 7(1), p.91.

Rosko, M.D. and Mutter, R.L., 2008. Stochastic frontier analysis of hospital inefficiency: a review of empirical issues and an assessment of robustness. *Medical Care Research and Review, 65*(2), pp.131-166.

Schlesinger, M., 2002. On values and democratic policy making: The deceptively fragile consensus around market-oriented medical care. *Journal of Health Politics, Policy and Law*, *27*(6), pp. 889-926.

Schreyögg, J., Stargardt, T., and Tiemann, O., 2011. Costs and quality of hospitals in different health care systems: A multi-level approach with propensity score matching. *Health Economics*, *20*(1), pp. 85-100.

Seiford, L. M., and Zhu, J., 2002. Modelling undesirable factors in efficiency evaluation. *European Journal of Operational Research*, *142*(1), pp. 16-20.

Shapiro, S., et al. September, 1960. Further Observations on Prematurity and Perinatal Mortality in a General Population and in the Population of a Prepaid Group Practice Medical Care Plan. *American Journal of Public Health* 50, pp. 1304–1317.

Shestalova, V., 2003. Sequential Malmquist indices of productivity growth: An application to OECD industrial activities. *Journal of Productivity Analysis*, *19*(2-3), pp. 211-226.

Shen, Z., T., Baležentis, X., Chen, and V., Valdmanis., 2018. Green Growth and Structural Change in Chinese Agricultural Sector During 1997–2014, *China Economic Review, 51*, pp.83-96.

Shephard, R. W. (1970). Theory of cost and production function. Princeton, NJ: Princeton University Press.

Sherman, H.D., and J., Zhu., 2006. Benchmarking with Quality-Adjusted DEA (Q-DEA) to Seek Lower Cost High-Quality Service: Evidence from a U.S. Bank Application, *Annals of Operations Research, 145*(1), pp.301-319.

Sloan, F. A., Picone, G. A., Taylor, D. H., and Chou, S. Y., 2001. Hospital ownership and cost and quality of care: Is there a dime’s worth of difference? *Journal of Health Economics*, *20*(1), pp. 1-21.

Siciliani, L., 2018. Hospital Economics: The Effect of Competition, Tariffs and Non-profit Status on Quality. *Health Econometrics* (pp. 263-283). Emerald Publishing Limited.

Street, A., 2003. How much confidence should we place in efficiency estimates? *Health Economics, 12*(11), pp.895-907.

Thomas, J.W. and Hofer, T.P., 1999. Accuracy of risk-adjusted mortality rate as a measure of hospital quality of care. *Medical Care*, pp.83-92.

Tipping, M.E. and Bishop, C.M., 1999. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61*(3), pp.611-622.

Tone, K., 2004. Malmquist productivity index. In W. Cooper, L. Seiford, and J. Zhu (Eds.), *Handbook on data envelopment analysis* (pp. 203-227). Boston, MA: Springer.

Tsai, T.C., Orav, E.J. and Jha, A.K., 2015. Patient satisfaction and quality of surgical care in US hospitals. *Annals of Surgery*, 261(1), p.2.

Tsoulfas, G. T., and Pappis, C. P., 2008. A model for supply chains environmental performance analysis and decision making. *Journal of Cleaner Production,* *16*(15), pp.1647-1657.

Yang, H. L., and Pollitt, M., 2009. Incorporating both undesirable outputs and uncontrollable variables into DEA: The performance of Chinese coal-fired power plants. *European Journal of Operational Research, 197*, pp.1095–1105.

Yip, W. and Hsiao, W., 2009. China's health care reform: A tentative assessment. *China economic review*, 20(4), pp.613-619.

Zaim, O., 2004. Measuring Environmental Performance of State Manufacturing through Changes in Pollution Intensities: A DEA Framework, *Ecological Economics*, *48*(1), pp.37–47.

CURRICULUM VITAE

Academic qualifications of the thesis author, Mr. HE Xinju:

- Received the degree of Bachelor of Economics from Southwestern University of Finance and Economics, June 2014.
- Received the degree of Master of Applied Economics from Hong Kong Baptist University, November 2015.

August 2019