An Empirical Study of New Zealand Secondary School Performance under the Qualification System of National Certificate of Education Achievement (NCEA)

Jane Qin

Master of Business (Mbus)

2016

An Empirical Study of New Zealand Secondary School Performance under the Qualification System of National Certificate of Education Achievement (NCEA)

Jane Qin

A thesis submitted to Auckland University of Technology in partial fulfillment of the requirements for the degree of Master of Business.

July 2016

Faculty of Business and Law

Primary Supervisor: Dr. Anil Narayan

Secondary Supervisor: Dr. Julie Harrison

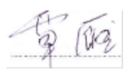
Table of Contents

List of Figures

List of Tables

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.



Acknowledgements

This thesis is a prolonged and exigent journey that required persistence and consistency. The voyage of expedition attributes its completion to the guidance, assistance and support of many people.

First, I would like to thank my two supervisors, Doctor Anil Narayan and Doctor Julie Harrison. Anil guided me from the outset in terms of developing the research proposal and conducting the literature review. I am grateful to Anil for his valuable assistance and feedback throughout the dissertation process. Julie has inspired me and cultivated my research endeavours. I am indebted to Julie for her continuous guidance, perseverance and always extending her helping-hand even during the school holidays. Anil and Julie, your patience and confidence in me encouraged me to do my best. Without your support, I would not have been able to complete this thesis.

I would like to acknowledge the help of Cyril Mako, Manager of Schooling Analysis Department at the Ministry of Education, for his assistance in school data access. Also, I am grateful to Andrea Raith for her technical support of the DEA software.

I would also like to thank all the administrative staff in the Postgraduate Research Office. They have provided me with professional advice and genuine support during my Master of Business study. In particular, Eathar Abdul-Ghani encouraged me to achieve what I believed I could; Noah Ghebremichael provided me prompt and professional advice whenever I needed. Thank you all.

Finally, without any implicating, a special thank you to my husband Dexter for his continuous support and limitless patience. He has been my sheet-anchor and companion throughout the journey.

Abstract

The National Certificate of Education Achievement (NCEA) was introduced as the major qualification system for New Zealand secondary schools in 2004. Given there is limited quantitative research that has examined the New Zealand public school performance under the NCEA system, this thesis attempts to investigate the major features of school performance and examine school productivity changes since the introduction of NCEA. Towards this end, a network data envelopment analysis (DEA) model was employed to measure school performance from multiple dimensions: the overall efficiency and its sub-efficiencies being cost efficiency, academic efficiency and academic effectiveness, of New Zealand high schools during two periods (2004-2006 and 2009-2011). Performance measures for all the sub-models in the network DEA model were selected incorporating the objectives of the primary stakeholder group (Government and the Ministry of Education).

New Zealand school performance under the NCEA system has similar features to that under the old qualification system in terms of decile effect, gender difference and integrated school performance. Significant improvements in school productivity in terms of overall efficiency, academic efficiency and academic effectiveness were found since the introduction of NCEA. Regression results suggest that low and medium decile schools and high Mäori density schools had overall a significant improvement in productivity with respect to academic efficiency, whilst co-educational schools on average enjoyed a significant increase in productivity for overall efficiency. These findings confirm that school overall productivity has increased since the introduction of NCEA, and the NCEA system has leveled the playing field for different student groups by narrowing the gap in performance between them.

The possible reasons for good/poor performance were also examined by looking at the Education Review Office (ERO) reports for the top ten and bottom ten schools. Good school practice during the later period (2009-2011) focused more on engaging student learning and the use of data, whereas in the earlier period (2004-2006), good practice related to teaching strategy and student support. These differences reflect the dynamic nature of the education system, to which the NCEA system, the national curriculum standards, as well as school performance measurement/management systems need to adapt.

Abbreviations

BoT	Board of Trustee
BSC	Balanced Score Card
CAAS	Consent to Assess Against Standards on the Directory of Assessment Standards Rules 2011
CIE	Cambridge International Examination
COLS	Corrected Ordinary Least Squares
CRO	Chief Review Officer
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
DRS	Decreasing Returns to Scale
ERO	Education Review Office
GAAP	Generally Accepted Accounting Practice
GCSE	General Certificate of Secondary Education
GDP	Gross Domestic Product
GNE	Gross National Expenditure
IB	International Baccalaureate
LPM	Linear Programming Model
IRS	Increasing Returns to Scale
MNA	Managing National Assessment
MoE	Ministry of Education
NAG	National Administration Guideline
NCEA	National Certificate of Education Achievement
NCEA L3 M&E	NCEA Level 3 Merit and Excellence
NEG	National Educational Guidelines
NPM	New Public Management
NZQA	New Zealand Qualifications Authority
NZQF	New Zealand Qualifications Framework
NZTC	New Zealand Teachers Council
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Squares
PISA	Programme for International Student Assessment
RTS	Returns to Scale
SES	Socio-Economic Status
SFA	Stochastic Frontier Analysis
SMART	Strategic Measurement and Reporting Technique
UE	University Entrance
UoLB	Use of Land and Buildings
VRS	Variable Returns to Scale

Chapter 1 Introduction

Internationally there has been a focus on the effectiveness and efficiency of educational expenditures in the public sector for at least the last thirty years (Ploom and Haldma, 2013). In line with this trend, New Zealand schools are expected to demonstrate continuous improvement in these areas, that is, each school is supposed to act in the best interests of the Government and other stakeholders, for example, parents and students, while making best use of the resources available. This raises the issue of how school performance can be measured to incorporate stakeholders' expectations and multiple educational objectives.

Given that there is no worldwide-accepted approach to the measurement of school performance (Bendikson, Hattie and Robinson 2011), this thesis attempts to develop a model to measure school performance by calculating school efficiency and effectiveness scores, and then analyses the change in school performance under the NCEA qualification system during the period between 2004 and 2011. In addition, statistics are produced for the top ten and bottom ten schools to reveal possible reasons for their performance based on their ERO reports.

The rest of this chapter explains the motivation and objectives of this study, the framework used to measure school performance and the structure of the thesis.

1.1 Motivation

Undoubtedly, increased school performance has a significant and positive influence on national economic performance (Hanushek, 2005), and as such, the New Zealand government has allocated over 6% of GDP to the expenditure on education each year since 2006 (OECD, 2013). As a consequence, New Zealand public school performance is subject to a high level of public interest, given that not only have the reforms to improve school quality consumed a large amount of resources but also involved a wide range of stakeholders. Nevertheless, what is considered important for creating school excellence is different for different stakeholders (Taut, 2008).

In the 1980s, a worldwide reform movement of education systems developed with the aim to introduce the concept of efficiency in school performance. Following this trend, since 1989, New Zealand has been at the forefront of the international school reforms by carrying out a series of structural changes known as "Tomorrow's Schools', with the focus on increasing school performance in terms of school efficiency and effectiveness (New Zealand Department of Education, 1988). Inspired by this origin, this thesis applies an economic model to measure New Zealand public school performance via the decomposition of overall efficiency into cost efficiency, academic efficiency and academic effectiveness.

There has been limited quantitative research addressing New Zealand school performance mainly due to three reasons. First, measuring school performance is a very complicated process owing to a variety of factors that influence school performance measurement. For example, different stakeholders have different views on what contributes to school excellence (Rivkin, Hanushek, & Kain, 2005), given that New Zealand public school performance varies along a number of dimensions such as teachers' salaries, school operational expenditures and academic outcomes. Second, the Government has continuously initiated educational reforms with an emphasis on equity and excellence in educational achievements. For instance, the introduction of NCEA and periodical updates on NCEA assessment standards. Finally, there are relatively limited quantitative data available to researchers due to privacy issues (Alexander, Haug and Jaforullah, 2010). Some publicly available data only reflect ratios, aggregated or qualitative information. For example, NCEA data are published in the form of ratio, school financial expenditure figures published on *education counts* [1] website are grouped into year levels and the ERO reports only display qualitative evaluation information for each school. As a result, measuring school performance and other elements associated with school performance is extremely complex and time-consuming.

Fortunately, there have been a handful of empirical studies that investigated New Zealand school performance (e.g., Harrison, 2008; Alexander et. al, 2010), and these papers provided referable experience in modelling New Zealand school performance. Yet, even though these studies measured school performance from multiple perspectives by including financial and non-financial, academic and non-academic measures, they were based on the data no later than 2001 due to their specific research purposes or the unavailability of desired data; thus their research results do not reflect the impact of NCEA on the New Zealand school performance (NCEA was introduced as the major qualification system for New Zealand secondary schools between 2002 and 2004).

[1] *Education counts* is the official website for the MoE to release statistics in relation to school and service.

Moreover, these studies employ a one-stage model to generate productive efficiency scores as the sole measure for school performance, which reflects the level of school managerial efficiency, rather than the effectiveness of achieving organisational objectives.

The introduction of NCEA represented a radical reform in the New Zealand school qualification system, as it was developed to enable more students to leave schools with qualifications. Thus, it replaced the old school qualification system by introducing more internal assessment standards to measure a wider range of competencies and skills (New Zealand Qualifications Authority, 2015). In the early years after the introduction, a number of principals and teachers could not understand the essence of the assessment standards and hence were unable to develop appropriate teaching strategies to engage student learning. As a consequence, NCEA was criticized for inconsistent internal assessment standards, poor planning and for encouraging mediocrity (NZ Herald, 2007). Against this background, several studies examined the influence of NCEA on New Zealand school performance or student achievement, such as Bendikson et al. (2011) and Shulruf, Hattie, and Tumen (2010). However, these studies measured school performance only using school or student NCEA outcomes. Thus there is a gap that little empirical research looks at school performance under the NCEA system using multiple perspectives of performance taken from the literature of management accounting, management control and production economics (Ploom and Haldma, 2013; Harrison, Rouse and De Villiers, 2012). This ultimately engendered the motivation of the author to examine New Zealand school performance, together with the productivity changes since the introduction of NCEA from financial and non-financial, academic and non-academic perspectives.

ERO reports are generally considered as the authoritative evaluation for school performance. However, ERO reports only provide qualitative assessment for school quality rather than quantitative results that can be compared or ranked. Moreover, ERO reviews for each school are generally not conducted on yearly basis. Therefore, to develop a performance measurement model for New Zealand schools that might help the Ministry of Education (MoE) identify schools with good or poor performance, especially in the years when the ERO reports are not available, creates another incentive for this research.

1.2 Research objectives

Given there has been limited quantitative research looking at New Zealand school performance from multiple perspectives under the NCEA system, this study models school performance from multiple dimensions and examines the productivity changes since the introduction of NCEA system. Three research objectives are reflected in the three research questions listed as follows:

RQ1.What components/elements should be included in a model of school performance that measures both efficiency and effectiveness?

RQ2. Have there been any productivity changes in school performance since the introduction of the NCEA qualification system? If so, what are the characteristics of these changes?

RQ3. What are the likely causes for the performance of the top ten and bottom ten schools based on their published ERO reports?

The above questions are designed to identify any distinctive features of New Zealand school performance under the NCEA system, together with any significant changes in school productivity. The possible causes found for good/poor school performance are expected to reflect the consistency between the results produced by the model and the ERO evaluation.

There are three expected contributions for this study. The first one is to provide an expanded performance measurement framework not only looking at school efficiency but also effectiveness; the second one is to establish a model that measures New Zealand public school performance via the decomposition of overall efficiency into cost efficiency, academic efficiency and academic effectiveness; the last one is to examine school performance and productivity changes under the NCEA qualification system to determine the impact of its introduction. In so doing, a method that might be of the MoE's interest, is developed to identify schools with good or poor performance. Moreover, the analysis on school performance and changes in school productivity under the NCEA system is expected to provide evidence to the MoE to evaluate the functions of NCEA for further decision-making.

1.3 Research methodology and method

This thesis employs a quantitative approach to model and analyse school performance supported by new public management (NPM) theory, stakeholder theory and Ramanathan's framework.

To help answer the research questions, first, a network data envelopment analysis (DEA) approach incorporating student socio-economic status (SES) is employed to produce comparative scores for school performance using NCEA academic outcomes as outputs and teachers' salaries and other operational expenditures as inputs. The overall school performance scores are supported by a decomposition of the overall efficiency in terms of cost efficiency, academic efficiency and academic effectiveness.

Second, based on the DEA efficiency scores, change in school productivity and its decomposition results are examined by calculating Malmquist indices to distinguish technical change (i.e., frontier shift due to the introduction of NCEA) and efficiency change (i.e., change in managerial efficiency). To identify the characteristics of these changes, regressions are undertaken to analyse the impact in terms of a number of exogenous factors being school ownership, co-educational status, Mäori density, decile level and school size.

Finally, statistics are produced to identify possible reasons for the performance of the top ten and bottom ten schools using information from their ERO reports. The comparison between the reasons for good and poor performance provides insights into what contributes to school excellence in the New Zealand context.

1.4 Structure of the thesis

This thesis contains five chapters that are summarised as follows:

Chapter 1 briefly introduces the key aspects of the research, including the motivation and research objectives.

Chapter 2 conducts a literature review concerning performance measurement for public secondary schools in New Zealand. It also characterises the DEA technique employed in this study.

Chapter 3 describes the research methodology and the methods of data collection and sample forming.

Chapter 4 presents and discusses the DEA results and Malmquist indices for New Zealand school performance and productivity changes. The analysis of the good and poor school performance is also provided.

Chapter 5 draws the conclusions and discusses the contribution and limitation of this study, followed by the areas for future research.

Chapter 2 Literature Review

2.1 Introduction

This chapter reviews the literature on school performance measurement in the New Zealand context. Section 2.1 introduces the chapter. Section 2.2 provides the background to the New Zealand secondary school system and the roles of stakeholders in the system. Section 2.3 reviews the literature on performance measurement of the public education system and informs the selection of an appropriate theoretical framework for public school performance measurement. Section 2.4 compares several performance measurement models in the education economic literature and explains why DEA is selected in this study.

2.2 Background

2.2.1 New Zealand school system

In New Zealand, a child at the age of five can enter a ‘new entrant’ class known as year 0 and the compulsory education commences when a child reaches the age of six and continues until sixteen (Guthrie & Tooley, 2007). Apart from the entry level, the New Zealand education system provides classes from year 1 to year 13. Specifically, primary schools deliver classes up to year 6 or 8; intermediate schools run classes for years 7 and 8 and secondary schools offer classes from year 9 to 13 or 7 to 13. Of note, some secondary schools provide continuing education to specific student groups, for example, teenage parents[2]. As a result, these schools serve years 14-15 students. However, a review of roll data for 2009-2011 suggested that the numbers of years 14-15 students in most schools are very small (less than 10).

The New Zealand secondary school system varies along a range of dimensions. Schools diverge as per the range of student ages for which they are tailored, the school location (urban/rural), ownership type (private/public) or whether they elect to be single-sex or co-educational.

In New Zealand, the traditional secondary schools offer classes to students between year 9 and year 13 whilst some secondary schools include year 7 to year 13 as a consequence of the amalgamation of traditional secondary schools with intermediate

[2] These schools are featured as Teen Parent Unit (Association of Teen Parent Educators New Zealand, 2011).

schools. In some rural locations, especially in those relatively inaccessible areas, schools may cover year 1 to year 13.

The types of school ownership are 'State-owned', 'Integrated' and 'Independent/private'. 'State-owned' is the most common type in the New Zealand context. State-owned schools are publicly funded and teaching in these schools is in accordance with the *New Zealand Curriculum*. Integrated schools are operated by private owners and are integrated into the public school system, but their land and buildings retain the status of being privately owned. Independent schools are those who elect not to join the state school system. However, they have to reach certain standards to get registered by the MoE. State-owned and integrated schools contribute about 94.8% of secondary school rolls (Education Counts, 2014b).

A number of New Zealand secondary schools with long history cater for only boys or girls while others are open to both. There are a large number of research findings suggesting that educational outcomes in single-sex schools are generally better than those in co-educational schools (Harrison 2008; Alexander et.al, 2010).

2.2.2 School accountability system in New Zealand

Generally speaking, accountability emphasizes the importance of explaining or justifying one party's conduct to another party as well as the consequences of good or poor performance (Wöbmann et al., 2007; Jackson, 1982). In New Zealand, public schools are accountable for their performance to the MoE and the ERO (Ministry of Education, 2014h; Education Review Office, 2014c). Schools with good performance are likely to be selected as benchmarks for schools suffering deteriorating performance, whilst those with poor performance inevitably capture the attention of the MoE and the ERO, followed by a series of interventions and more frequent evaluation reviews.

As the government agencies such as the MoE and the ERO play important roles in school performance evaluation, it is essential to be conversant with the New Zealand school accountability system before modelling its performance measurement system.

The New Zealand education system is overseen by a number of separate entities known as the MoE, the ERO, the New Zealand Qualifications Authority (NZQA), the New Zealand Teachers Council (NZTC) and individual schools governed by schools' Board of Trustees (BoT). Each of them has specific responsibilities and accountabilities (Ministry of Education, 2014l).

The MoE acts on behalf of the Government in the educational system with respect to policy making and funding approval (Parliamentary Counsel Office, 2014). In addition, the MoE provides directives to the education agencies such as the NZQA, the ERO and the NZTC, as well as the education providers (primary/secondary schools and tertiary institutes) in New Zealand. The MoE is accountable to the Minister of Education.

The ERO audits school performance as per relevant statutes and publishes periodic evaluation reports for each school. The Chief Review Officer of the ERO is appointed by the Minister of Education (Education Review Office, 2014a). The ERO is accountable to the Minister of Education.

The NZQA oversees academic qualifications for secondary school students under the NCEA system. The Authority's board is appointed by the Minister of Education, and it is responsible for shaping the direction for NZQA to function in line with the Government's goals. The NZQA is accountable to the Crown via the Minister of Education.

The NZTC is in charge of managing the registration, de-registration and renewal of the national certificates for New Zealand teachers. It is accountable to the Minister of Education.

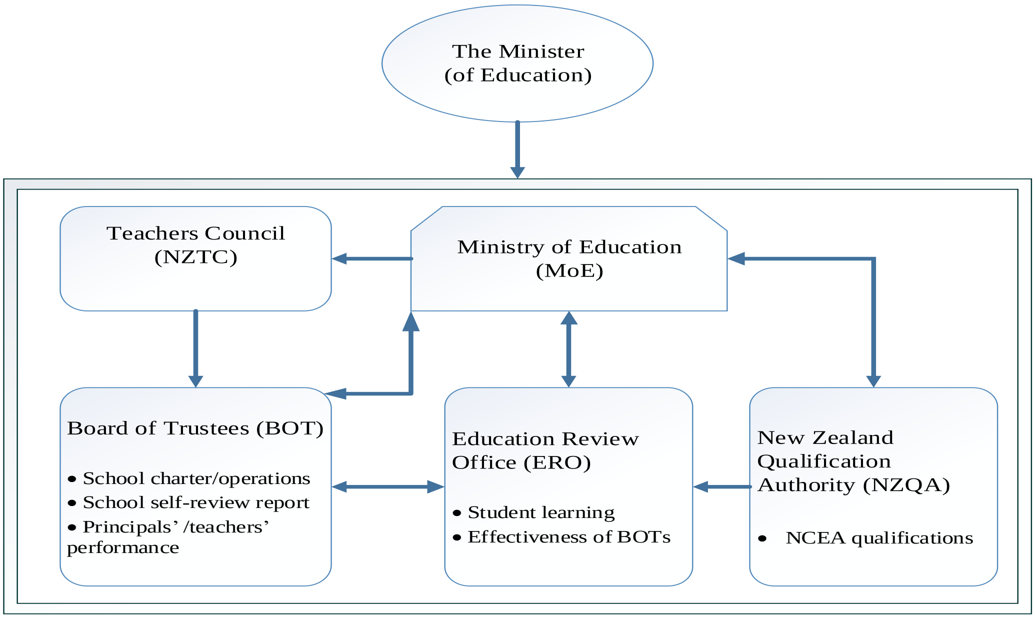
A school's Board of Trustees (BoT) is legally responsible for the governance of the school. It appoints the school principal who has operational control over the school. Hence, a school principal is accountable to the school's BoT that is accountable to the MoE and the ERO.

Since the ERO, the NZQA and the NZTC are intimately related to the MoE through the Minister and BoTs are accountable to the MoE, it is reasonable to conclude that the MoE plays a leading role on behalf of the Government in the New Zealand education sector and therefore the Government and the MoE form the major stakeholder group in the New Zealand education system.

Figure 2.1 illustrates the school accountability system in New Zealand.

To elaborate more detailed functions of the major components in Figure 2.1, subsection 2.2.2.1 discusses the role of the MoE; 2.2.2.2 describes the NCEA system; 2.2.2.3 explains what the ERO does and 2.2.2.4 demonstrates New Zealand school accountability.

Figure 2.1 School's accountability system in New Zealand



Adapted from Malik (2011, p.15)

2.2.2.1 The role of the MoE

The MoE supplies funding for state-owned and integrated schools and develops policy guidelines for each BoT to improve school performance in alignment with the national educational goals, known as National Educational Guidelines (NEGs), which consist of the guidelines for the school performance management system (Ministry of Education, 2014e), the guidelines for teachers' performance measurement and the guidelines for measuring principals' performance (Ministry of Education, 2014f).

NEGs contain five components: national educational goals, foundation curriculum policy statements, national curriculum statements, national standards and national administration guidelines (NAGs) (Ministry of Education, 2014a).

According to the MoE's "Statement of Intent 2011/12-2016/17" which outlines the objectives to be achieved in the six years, the MoE has four objectives related to secondary education:

- Strengthen literacy and numeracy teaching and learning;
- Maximize the number of school-leavers with qualifications;
- Reinforce the support to Mäori learners' success;
- Deliver core functions to support students' engagement and success in education. (Ministry of Education, 2011)

The ultimate goal is to "Lift aspiration and raise educational achievement for every New Zealander" (Ministry of Education, 2014c). Towards this end, the MoE plays a role as "the Government's lead advisor on the education system, shaping direction for education agencies and providers and contributing to the Government's goals for education" (Ministry of Education, 2014d). The MoE's functions are:

- Working out education policies and originating initiatives to ensure the implementation of these policies;
- Supervising policy implementation and evaluating the subsequent outcomes;
- Supplying funds for school operations/educational resources;
- Paying teachers' salaries and engaging influential stakeholders and professional contractors to achieve desirable educational outcomes.
  (Ministry of Education, 2014d)

In short, the role of MoE is reflected in two aspects that are particularly relevant to secondary schools: educational policy making and funding for schools. To maximize the effectiveness of the policies and the efficiency of the funding for schools, the MoE takes steps to monitor the effect of policies on school performance and aid failing schools.

The MoE supervises the outcomes of policies by analysing the data collected from secondary schools[3] such as retention of students and number of school leavers with NCEA Level 2 or above. In the circumstance that a deteriorating school is identified, the MoE may make informal/formal interventions to assist the school to improve performance. The types of interventions are subject to Section 78 of Education Act 1989 (Ministry of Education, 2014h).

The MoE supplies funds to all state-owned and integrated schools in New Zealand. The types of funds are categorized as operational funding, school staffing, additional payments to schools and payments to individuals (Ministry of Education, 2014i). Payments to teachers and other staff/individuals are transferred directly by the MoE to an employer or employees as per the relevant terms of their employment contacts (Ministry of Education, 2014j).

[3] The MoE collects data in terms of eight indicators: "retention of students in senior secondary schools, truancy, stand-downs and suspensions, exclusions and expulsions from school, early school leaving exemptions, school leavers with no qualifications, school leavers with NCEA Level 2 or above, and school leavers with a university entrance standard" (Ministry of Education, 2014g).

The operational funding rate for each school's operations is determined by a number of factors such as school roll, school type, school decile level and specific needs of students (Ministry of Education, 2014k). Even though the funds supplied by the MoE are supposed to cover the required expenses associated with school operation, most schools raise additional funding mainly from non-governmental sources such as international students (if applicable), parents' donations, school trading and the land/building use grants (NZ Herald, 2009). In general, the fund-raising from non-governmental sources at high decile schools (decile>6) is easier than low decile schools (decile <4) (Malik, 2011).

2.2.2.2 NCEA qualification system

NCEA was introduced as the major qualification system for New Zealand secondary schools in 2002 and was fully entrenched by the end of 2004. It replaced the old qualification system comprising School Certificate, University Entrance, Sixth Form Certificate and University Bursary qualifications (New Zealand Qualifications Authority, 2015).

Compared with the old qualification system, the NCEA system provides a wider range of standards to reflect students' competencies and skills, in particular, more standards are issued for internal assessments. Since the introduction of NCEA, more school leavers have obtained qualifications which are accepted by most employers and universities in New Zealand and overseas (New Zealand Qualifications Authority, 2015). This outcome was expected by the MoE as it has been carrying out a series of reforms to accomplish equity and excellence in educational achievements. In this sense, the NCEA system is designed to reflect student learnings in both academic and vocational spheres.

The process of achieving the NCEA qualifications for secondary school students is administered by the NZQA.

The NCEA certificates are issued to a student by calculating how many credits he/she has been awarded according to the standard he/she meets in the subject/ programme he/she has selected. The standards are categorized as unit standards and achievement standards. Unit standards are made to internally assess a student's performance by "pass" or "fail", while achievement standards are set for both internal and external assessments grading as "achieved, "merit" or "excellence". A student can obtain the credits for one course as long as he/she passes the unit standards or at least receives "achieved" under achievement standards (New Zealand Qualifications Authority, 2013).

There are three levels of NCEA qualifications: NCEA Level 1, NCEA Level 2, and NCEA Level 3. The higher the level, the more complex the requirements are. In general, year 11 students obtain Level 1 qualification; year 12 students obtain Level 2 qualification and year 13 students obtain Level 3 qualification. The achievement level for Level 1 qualification is 80 credits in any level including 8 credits in literacy and 8 credits in numeracy; the achievement level for Level 2 qualification is 80 credits with 60 credits out of the 80 credits from Level 2 courses; the achievement level for Level 3 qualification is 80 credits with 60 credits from Level 3 papers and the remaining from Level 2 or 3 courses (New Zealand Qualifications Authority, 2013).

Where the requirements for Level 3 qualification are satisfied, a student can receive NCEA with Excellence where at least 50 of the 80 credits are ranked as Excellence; similarly, to achieve NCEA qualifications with Merit, a student must gain at least 50 of the 80 credits ranked as Merit.

A student can attain University Entrance (UE) if he/she achieves 60 credits or above at NCEA Level 3 (New Zealand Qualifications Authority, 2013).

2.2.2.3 The role of the ERO

The ERO is a government department that was formed subject to the State Sector Act 1988. The head of the ERO is the Chief Review Officer (CRO). The major job of the CRO is to organize reviews of performance for pre-tertiary education providers such as schools or early childhood services under Part 28 of the Education Act 1989 (Education Review Office, 2014b).

The ERO evaluates "the quality of education and care in schools and services, and the implementation of Government education priorities" (Education Review Office, 2014b, p. 5). The ERO is independent from schools, the MoE, the NZQA and other agencies. This allows the ERO to give the Government "assurance of the quality of education" (Education Review Office, 2014b, p. 5).

The ERO provides as a "guardianship and improvement" role in the education sector and its independence from schools and other agencies enables it to produce authoritative evaluation on school performance and policies made by the MoE. Therefore, the ERO plays both accountability and improvement role in the education sector. In addition, the ERO also advises the Minister about the formulation of future priorities and reports the emerging issues identified in the performance review process (Education Review Office, 2014b, p. 5).

The education reviews concentrate on student learning and how school strategies, policies and practices help improve “student engagement, progress and achievement” (Education Review Office, 2011, p. 2). To conduct a performance review for a school, a review officer works with the school and identifies the strengths to be carried forward and the weaknesses to be addressed. The ultimate goal of the review is to enable schools to provide quality educational service by maximizing the efficiency of the funds supplied by the Government and other sources.

The ERO’s review concerns “school curriculum, school self-review, national evaluation topics, the Board Assurance Statement and student health and safety” (Education Review Office, 2014c, p.5). The ERO’s evaluation is based on six dimensions: student achievement, governance, leadership, teaching, school culture and engagement with parents/families (Education Review Office, 2014c, p.28).

The ERO reviews for a school are generally undertaken every three years. However, they can be performed in shorter intervals where the ERO has concerns with the school’s deteriorating performance (Education Review Office, 2014c, p.31-32).

It is worth mentioning that ERO reports are qualitative with limited or no quantitative analysis of school performance. In this research, ERO reports are analysed to produce possible reasons for good or poor school performance (see chapter 4).

2.2.2.4 Board of Trustees (BoT) and school accountability

The BoT is a statutory body created under section 93 of the Education Act 1989. A BoT is elected and consists of parent representatives, the school principle, teacher representative and student representative, who ‘reflect the ethnic and socio-economic composition of the school’s student body” (Ministry of Education, 1989). The BoT appoints the school principal who has operational control over the school.

A school’s BoT is vested in the governance of the school by the Education Act 1989 and is responsible for formulating policies for school management, developing a performance evaluation scheme to assess the performance of teachers and principals, preparing school budgets and financial reports, and conducting self-reviews for the school's performance in accordance with the MoE's and the ERO's requirements (Ministry of Education, 2015; Smelt, 1998).

The responsibilities of the school's BoT reflect the school's accountability to the MoE and the ERO. In brief, school's accountability includes:

- Formulate a school charter for school operation and management (MoE);
- Provide a self-review report (MoE and ERO);
- Establish and implement a performance management program for school principal and teachers (MoE);
- Provide any financial and non-financial information as required (MoE).

In summary, the MoE and other Crown entities such as the NZQA and the ERO currently evaluate/measure different facets of school performance. For each school the NZQA evaluates and records the academic outcomes of secondary students and the ERO accesses the quality of education provided by each secondary school in terms of the governmental regulations, the efficiency of the state funding, the ownership interests and other aspects in relation to student learning. The NZQA and the ERO share the evaluation information with the MoE to help it develop new policies to improve existing strategies. For each school the BoT is accountable to the MoE and the ERO for school performance in terms of school governance, self-review, information reports and performance management for the school principal and teachers.

While a number of theoretical perspectives support that academic outcomes can be considered as the major measure for school performance, the review above on the New Zealand secondary school system shows that academic outcomes cannot reflect all aspects of performance for each secondary school. Moreover, while ERO reports provide a broader perspective, they only present qualitative evaluation results that cannot be compared and ranked. These factors support that for comparison, school performance measured quantitatively should incorporate multiple theoretical perspectives taken from management accounting, management control and production economics.

Section 2.3 reviews the literature on performance measurement in the public sector and discusses the key features of how to develop a quantitative model to measure public school performance.

2.3 School performance measurement

This section answers research question one, that is, "what components/elements should be included in a model of school performance that measures both efficiency and effectiveness?" (see section 1.2).

According to Rouse and Putterill (2003, p 795), performance measurement can be defined as "the comparison of results against expectations with the implied objective of learning to do better". Surrounding this definition, the most popular issues that are relevant to modelling a performance measurement system include:

Q1. Whose performance is observed?

Q2. What expectations are there for the performance and how can these be transferred into measurable criteria?

Q3. What attributes are to be measured and what are their relationships to the organizational environment? (Rouse and Putterill, 2003)

The answer to the first question is straight forward in this study, as looking at New Zealand public school performance has been clearly stated in the research objectives. However, compared to question one, question two is much more complicated, given that the New Zealand education sector is part of the public sector, which is characterised by multiple stakeholders and multiple strategic objectives. Therefore, understanding the key issues on performance measurement in the public sector is a starting point to develop a performance measurement system for New Zealand public schools. In this regard, section 2.3.1 reviews the literature on performance measurement issues in the public sector; section 2.3.2 discusses a performance measurement framework for the New Zealand public schools. These two sections attempt to address the issues raised in question two. Section 2.3.3 is expected to answer question three, which describes the selection of both controllable and environmental variables in the performance measurement framework established in section 2.3.2. At the end of section 2.3.3, research question one will be served.

2.3.1 Theoretical perspectives on performance measurement in public sector

The public sector has been under continuous pressure to improve its performance since the early 1980s because the public believes that governments are excessively large, inefficient, ineffective and unresponsive to change (Pollitt and Summa, 1997). As a result, the "new public management" (NPM) philosophy emerged, which advocates that the public sector can improve its performance by exercising private sector management techniques to increase efficiency (best value of money) and effectiveness (strategic achievements) (Hood, 2000; Lapsley, 2008; Ploom and Haldma, 2013). In other words, it is argued that a public sector organisation should develop its own performance measurement system for the provision of public services, and this system should be subject to adjustments for relevant changes in business environment and organizational structure so as to retain its relevance and usefulness (Bititci, Turner and Begemann, 2000).

However, Ittner and Larcker (1998, p.233) questions "whether private sector notions of performance measurement and accountability are applicable in the public sector" and Fryer, Antony and Ogden (2009, p.491) points out "expected improvements in public sector performance have not yet materialised". Even though a performance measurement system is considered to be helpful with the achievement of organizational objectives and stakeholders' prospects as long as it is designed to reflect the organization's objectives, there may be a conflict between the goals of the organization and the stakeholders (Brudan, 2010). Moreover, as Näsi (1995) demonstrates, an organization in the public sector may actually have no objectives of its own but regard its major stakeholders' demands/interests as the organizational goals. Therefore, it can be argued that the objectives of public organizations represent the expectations of their stakeholders (Ploom and Haldma, 2013). However, there could be contradictions between the goals of different stakeholders and this raises the issue of how to design a performance measurement system that links the performance to organizational goals in considering the major stakeholders' expectations, because an organization can only survive if it satisfies the needs of its main stakeholders (Calton, 1993). So far as this is concerned, stakeholder theory provides a theoretical support to the design of performance measurement system as a means to optimize the process of organizational performance management (Jones, 1995; Freeman, 1984).

Stakeholder theory emphasizes managing stakeholders to achieve organizational objectives. There are two branches in the literature: instrumental and normative theories. Instrumental theory highlights "mutual trust and co-operation" that can produce superior solutions whereas the normative stakeholder theory refers to the reasons for "promoting stakeholder interests" irrespective of the existence of the interests (Alam, 2006). Both approaches centre on "who is a stakeholder", "what stakes they pursue" and "how management can undertake strategies to prioritize these stakeholders" (Alam, 2006, pp. 210-211). In the education sector, there could be multiple stakeholders that have significant influence over the formulation of organizational objectives. According to stakeholder theory, identifying the stakeholders should be done prior to ascertaining the organizational objectives for the design of a performance measurement system.

Given that it is unrealistic for an organization to satisfy its stakeholders equally, there are a number of attempts made to help organizations identify stakeholders and prioritise the demands from different stakeholders. Typically, Clarkson (1995) categorized stakeholders as primary and secondary stakeholders. Primary stakeholders are ranked higher as their support is crucial to the organization. For example, primary stakeholders could be shareholders, government, customers, suppliers, employees or creditors. In contrast, secondary stakeholders are not regarded as being vital to the survival of the organization. They could include environmental protectors or social workers. Mitchell, Agle and Wood (1997) attempted to create a model to prioritize stakeholders based on a series of crucial dimensions being "power, legitimacy, and urgency perspectives". As such stakeholders can be ranked from least to most important in terms of eight categories being "non-stakeholder, dormant, discretionary, demanding, dominant, dangerous, dependent and definitive stakeholders" (Alam, 2006, p. 212). As for the relationship between stakeholders and organizational objectives, Atkinson and McCrindell (1997) points out that a public sector organization's primary objective is shaped by legislation or the priorities of the Government, and the secondary objectives may reflect the expectations from a variety of stakeholders.

In the education sector, it is reasonable to include students, teachers, parents, the Government and the general public as the stakeholders based on the approach proposed by Atkinson and McCrindell (1997). Further applying this approach to the New Zealand education system, the stakeholder group in charge of formulating the primary objective is the Government and its agencies (i.e., the MoE, the ERO and NZQA etc.). For example, the MoE has the power of policy setting and school funding under the legislation. Therefore, the Government together with its agencies comprise the primary stakeholder group according to Clarkson (1995), or a "definitive stakeholder" group as per Mitchell et.al. (1997), given that the support from the Government and its agencies is crucial to the New Zealand education system thus this group has significant "power, legitimacy and urgency". Accordingly, to develop an effective performance measurement system, measures must be designed to reflect the expectations and goals of the Government.

However, even though the primary/definitive stakeholder is identified and subsequently the organizational objectives could be defined in alignment with the primary stakeholder's expectation, the conflicts between the objectives of different stakeholders may exist and they could become the obstacles to establishing a performance measurement system. Brignall and Modell (2000) generalizes three key groups of stakeholders that may significantly influence an organization in the public sector: "funding bodies, professional groups within provider organization and purchasers". Funding bodies concentrate on "financial results and resource utilization", professional groups emphasize "quality and innovation", and purchasers are concerned with "quality, resource utilization and competitiveness" (Brignall and Modell, 2000, p.291). The relative power of each key group of stakeholders and the interplay between them in the organization dominate the "balance and integration" of performance measures in a performance measurement system (Ittner and Larcker, 1998). The concept of "balance" refers to the different measures in the model needed to reflect the interests of various stakeholder groups whereas "integration" means that these measures are integrated and causally connected with organizational performance outcomes (Brignall and Modell, 2000).

With regard to the New Zealand education system, the MoE is the funding body on behalf of the Government and therefore it has the highest level of influence over the design and collection of performance measures. In reality, the MoE takes the responsibility of distributing the Government grants to each state-owned or integrated school within the budgetary constraints and monitors the quality of education provided by each school through a range of mechanisms such as school periodical reporting and administration of local branches. Therefore, the MoE's goals reflect the Government's expectation for school success: maximizing educational outcomes of student learning within the constraints of the budgetary level of resources.

In the New Zealand education sector, the direct purchasers of education service are students and their parents. This group focuses on achieving qualifications for job seeking or university entry, and they pay less attention to school expenditures than the MoE and the ERO. The indirect purchasers/consumers could be the general public or taxpayers, the community and the Government, who emphasize improving the quality of workforce and efficiency of resource utilization in order to ensure a constant growth of public wellbeing and national economy.

The professional group generally includes teachers, principals, non-academic employees and BoTs. This group are concerned with the brand of the school, the environment for student learning, working conditions and performance management system for employees.

Further applying Brignall's approach to the design of performance measurement system in the New Zealand public education system, the measures must be balanced to sufficiently reflect the objectives of the Government and the MoE (primary/definitive stakeholder), and simultaneously attend to the expectations of other stakeholder groups, for example, students and parents. Moreover, these measures must be integrated and linked to performance drivers/indicators so as to reflect the causal relationships between the measures and performance (Kaplan and Norton, 1996a).

In summary, according to the doctrine of NPM, the performance of public sector should be improved by increasing the public wellbeing and national wealth without the waste of resources (Brignall and Modell, 2000). Therefore, performance measurement systems for the public sector should be designed to reflect the efficiency of using the resources, as well the effectiveness of achieving the organizational objectives (Lockheed and Hanushek, 1994). Accordingly, a performance measurement system is expected to be designed on the basis of unambiguous standards and measures, the low consumption of resources, as well as improved accountability which allows the measures to be adhered to (Bititci, Carrie and McDevitt, 1997). In this regard, the selection of performance measures is a complex process which incorporates a wide range of theoretical concepts and techniques to address the issues in relation to stakeholder's satisfaction, performance norms and organisational environment (Greiling, 2006).

Based on the above discussion, the next section identifies and develops a suitable performance measurement framework for New Zealand public schools.

2.3.2 Performance measurement framework

In general, a framework is considered as a useful way of "thinking about systems for modelling purposes" (Rouse and Putterill, 2003, p.791). Herein, a system refers to "an assemblage of entities observed as acting cohesively" (Beer, 1972). To model a school performance measurement system, it is essential to develop a framework to assist in the process by identifying boundaries, defining dimensions and identifying the relationships among these dimensions (Rouse and Putterill, 2003). In this sense, an effective framework can help determine measurable criteria based on the expectations for school performance and provide a means to transfer these expectations into the criteria (see Q2 in section 2.3).

Thus, before selecting a suitable performance measurement framework for New Zealand public schools, it is essential to determine what is expected for school performance. As mentioned in the previous section, performance of education systems is expected to be measured in terms of educational achievements along with the best use of resources (Brignall and Modell, 2000). Performance can therefore be improved by ensuring greater efficiency and effectiveness. Lockheed and Hanushek (1994, pp. 5-6) explains these two concepts in a framework regarding the performance of education systems:

"*A more efficient system obtains more output for a given set of resource inputs, or achieves comparable levels of output for fewer inputs, other things being equal. [...] Educational effectiveness is whether or not a specific set of resources has a positive effect on achievement and, if so, how large this effect is*."

Hence, efficiency refers to the use of resources whilst effectiveness reflects how well strategic objectives were achieved in the education process (Ploom and Haldma, 2013). Accordingly, a framework for New Zealand school performance measurement is expected to demonstrate how to identify and define measureable criteria in New Zealand educational context.

According to the literature on organisational performance measurement, most of the early frameworks were designed for the private sector to help increase organizational accountability by linking strategy and performance to multiple stakeholder perspectives, and incorporating financial and non-financial measures within the framework (Harrison, Rouse and De Villiers, 2012). For example, the strategic measurement and reporting technique (SMART) pyramid attempted to make the above links (Cross and Lynch, 1989) and the balanced score card (BSC) incorporated financial and non-financial measures to evaluate organizational performance (Kaplan and Norton, 1992, 1996a). These approaches were initially developed for the private sector where stakeholder objectives were generally linked to financial performance. However, they could be useful in the public sector where the accountability of an organization is associated more with non-financial performance, and the importance of the links between strategic objectives and performance measures encapsulated in these frameworks is equally applicable. In essence, performance measurement systems for the private sector assist managers in decision making and management control, whilst those for public sectors aid public organizations in improving public accountability, efficiency of resource utilization and effectiveness of service provided (Brignall and Modell, 2000; Harrison et. al, 2012).

The BSC is one of the most popular frameworks adopted in the public sector to measure and manage organizational performance (Rigby and Bilodeau, 2009). The BSC measures organizational performance through four perspectives that relate to customers, internal process, organizational innovation and learning, and financial performance. Different measures representing organizational objectives are chosen and balanced for each perspective in terms of the nature and objective of the measures. As such, the BSC becomes a tool of “translating strategy into action” by connecting strategy with performance norms (Kaplan and Norton, 1992). Moreover, it provides a mechanism for managers to ensure that the improvement in one area is not obtained at the cost of another. However, even though the BSC focuses on the expectations of the shareholders/owners/government and customers/purchasers, neither does it specify the dimensions of performance reflecting success, nor provide effective method on how to balance and weight the measures along different dimensions. Therefore, it is less applicable in the case of making comparative assessments among schools (Harrison, 2008; Lingle and Schiemann, 1996; Kenny, 2003; Atkinson, Waterhouse and Wells, 1997).

Compared to SMART and the BSC, Ramanathan’s framework is more relevant in terms of measuring efficiency and effectiveness for public organisations. The major advantage of this framework is that it incorporates both management control and performance measurement perspectives (Harrison et.al, 2012). Ramanathan’s framework was developed to decompose benefit and cost around inputs, outputs and outcomes (Ramanathan, 1985). Herein, inputs are resources as the inflows of the production process while outputs are achievements as the outflows. The amount of outputs may cause changes in outcomes which can be regarded as states at specific points in time (Rouse and Putterill, 2003). Efficiency is measured by how much outputs are gained per unit of inputs; effectiveness is reflected by how much outcomes are achieved per unit of outputs.

Equation 2.1 illustrates how Ramanathan's framework decomposes the benefit-cost criteria into four interrelated control links (see more details in Appendix I).

Equation 2.1 Decomposition into control links for benefit-cost criteria

$$\frac{B}{C} = \left(\frac{B}{OC}\right) \times \left(\frac{OC}{O}\right) \times \left(\frac{O}{I}\right) \times \left(\frac{I}{C}\right)$$

Each ratio in the equation represents a control linkage associated with the social benefit per unit of cost (*B/OC*), which is expressed by the multiple of social benefit per unit of outcomes (*B/OC*), outcomes per unit of outputs (*OC/O*), outputs per unit of inputs (*O/I*) and inputs per unit of cost (*I/C*). If this framework is applied to measure performance for a production process, *I/C* measures cost efficiency, *O/I* reflects productivity efficiency and *OC/O* represents effectiveness.

According to Ramanathan (1985), the whole set or subset of these control linkages could be used to evaluate the performance of any part of the organization. In other words, to measure performance incorporating stakeholders' objectives, the control linkages in Ramanathan's framework should be chosen as variables on which the organization directly controls. In this sense, from the view of the Government and its key agency (i.e. the MoE), organizational performance should be evaluated through all the linkages in the model. In contrast, an organization is mainly concerned with outcomes, outputs, inputs and costs, while the consumers/users look at outcomes and outputs. Applying this framework, Harrison (2008) produced a number of control linkages for a typical education system, from which Table 2.1 is adapted to demonstrate the updated linkages under the NCEA system.

Table 2.1 Application of Ramanathan's Framework to Education System

Control Linkage	Education System
Social mission	Enable students to achieve full potential in society and contribute to their and New Zealand's future.
Social benefits	Increase in lifetime income of workforce, increase in economic growth, reduction in private training, reduction in unemployment benefits
Outcomes	Number of school leavers with NCEA Level 2 or above, number of school leavers with university entrance, number of students staying on in education to age 17, number of 19-year-olds who have attained a Level 2 or higher New Zealand Qualifications Framework (NZQF) qualification, number of students with NCEA Level 3 Merit or Excellence endorsements
Outputs	Number of students funded, number of examination passes, number of higher grades achieved in examinations, number of hours of student learning, average number of years in school, number of days of truancy, number of sporting activities or achievements, number of cultural activities, number of social activities, and number of artistic activities, number of Mäori and Pasifika passes
Inputs	Teacher numbers, administration staff numbers, number of registered Mäori and Pasifika school teachers, classroom equipment hours, number of classrooms, number of books, computers and other educational materials, innate ability and knowledge of students, students' willingness to learn
Costs	Teachers' salaries, management and administrators' salaries, other staff salaries, expenditure on learning resources, expenditure on property and maintenance, all other school expenditure

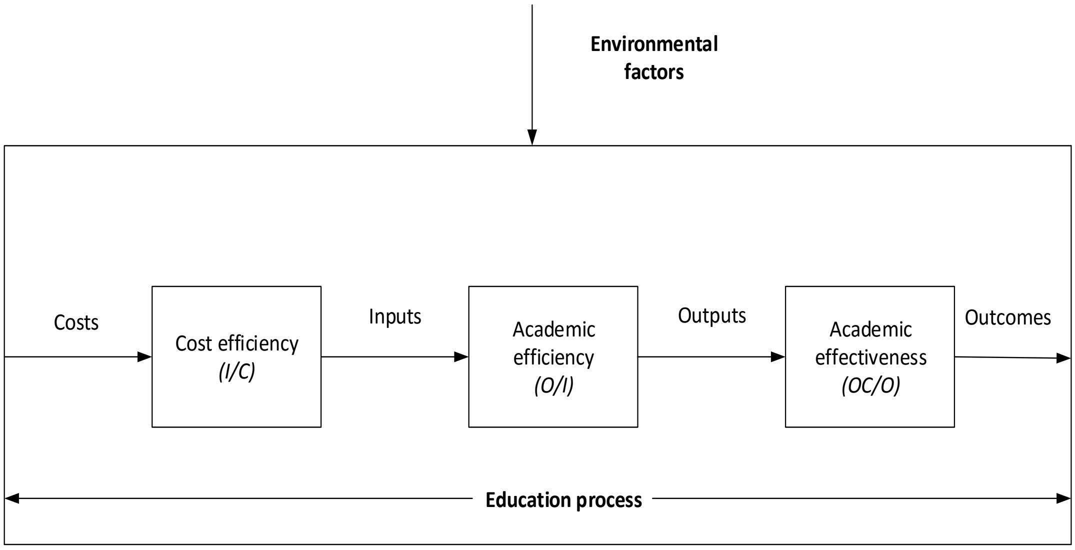
Adapted from Harrison (2008, p. 27)

Given that only the linkages relating to costs, inputs, outputs and outcomes are under the control of an individual school (Table 2.1), thus, only the linkages of *I/C*, *O/I* and *OC/O* in Equation 2.1 are relevant when constructing a performance measurement framework for public schools. Because an education process is also a production process which transfers resources/costs into outcomes (Rouse and Putterill, 2003), the overall efficiency of the process can be reflected by *OC/C*, which can be decomposed into the elements of *I/C*, *O/I* and *OC/O* following the reverse flow[4] of control linkages defined in Equation 2.1. As such, a proposed framework for public school performance measurement has been developed as shown in Figure 2.2, where the elements representing *I/C*, *O/I* and *OC/O* are renamed as cost efficiency, academic efficiency and academic effectiveness in the educational context. These efficiencies are expected to reflect school performance from different facets. Of note, environmental factors related

[4] Since Equation 2.1 is a decomposition process from social benefits to costs, whilst a productivity process begins with costs and ends up with outcomes. Therefore, the linkages in a productive process are connected in the reverse order of that shown in Equation 2.1.

to the education system are included in the framework, given that a school performance measurement system is supposed to reflect the change in school environmental context (Bititci, Turner and Begemann, 2000).

Figure 2.2 Proposed framework for school performance measurement



With regards to the selection of measures (i.e., inputs and outputs) for each efficiency element[5] in the framework, stakeholders' expectations are expected to be linked to these measures. Ideally, the goals of all the stakeholder groups can be satisfied in a generic framework where all the measures are set properly. Yet, this requires the inclusion of a large number of performance measures in the model (shown in Table 2.1), which inevitably leads to the problem of "proliferation of measures" (Harrison et. al, 2012, p. 246). To relieve this problem, the measures selected must be able to reflect the "shared objectives" in the organisation, or the expectations of the primary stakeholder(s) where different stakeholders have conflicting objectives against each other. Therefore, in the New Zealand educational context, measures for school performance in the framework are expected to represent the primary stakeholders' expectations (i.e., the Government and the MoE), given that different views exist among stakeholders on what contributes to good school performance in New Zealand.

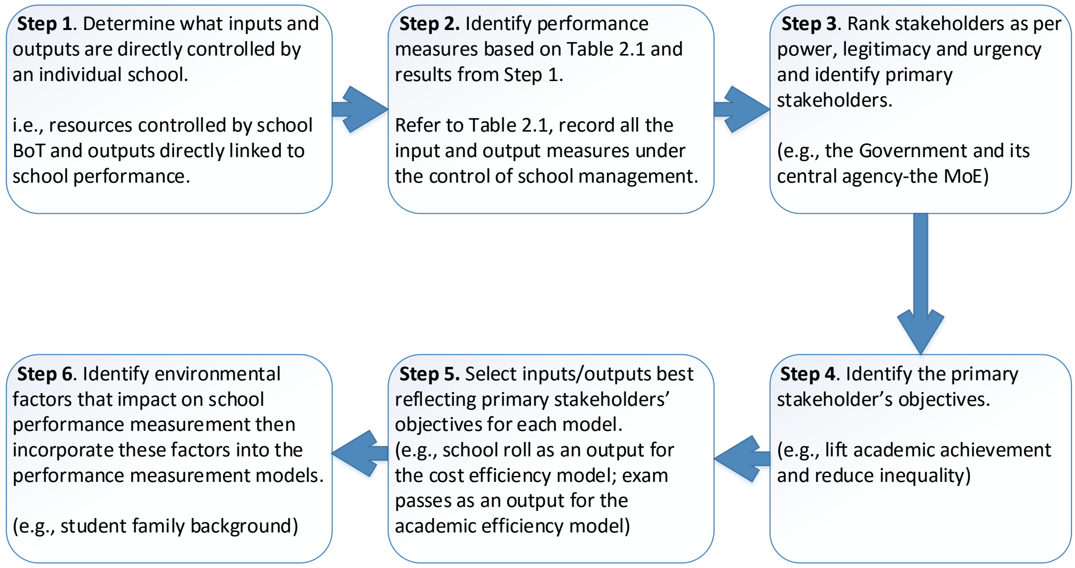
Thus, to refine the proposed framework in Figure 2.2, six steps listed below have been developed based on the preceding discussion and applied to each of the elements for cost efficiency, academic efficiency and academic effectiveness (see also Figure 2.3):

[5] These terms refer to cost efficiency, academic efficiency or academic effectiveness in the framework.

- In Step 1, inputs and outputs directly controlled by individual schools are identified in terms of what resources are controlled by school management and what outputs are directly linked to school performance. For example, as a result of "Tomorrow's Schools", school expenditure is distributed from the MoE, but the responsibility for the use of that money is decentralised to school level (i.e., BoT) (Fiske and Ladd, 2000). Therefore, school expenditure can be recorded as under the direct control of school management.
- In Step 2, all the proposed inputs and outputs available for the collection are recorded in reference to Table 2.1, such as teachers' salaries, exam passes and operational expenditures.
- In Step 3, primary stakeholders are identified who have the highest power, legitimacy and urgency. For instance, the Government and the MoE form the primary stakeholder group in the New Zealand education sector.
- In Step 4, the primary stakeholders' objectives are identified. For example, the MoE's objective is to "lift academic achievement and reduce inequality" (Ministry of Education, 2014c).
- In Step 5, input and output measures are linked to the primary stakeholders' objectives and only those inputs and outputs best reflecting the primary stakeholders' objectives are selected to measure school performance, given that in public sector, the objectives of primary stakeholders generally reflect the intention of the major part of the community. In this regard, to measure New Zealand school performance, the school roll can be chosen as an output for cost efficiency model as it is in line with the intention of 'equality', because school roll measures the number of students funded irrespective of the differences in student ethnic groups, gender and school deciles. The number of exam passes is an ideal output for academic efficiency model, given that it reflects school academic achievement.
- Finally, environmental factors that cannot be controlled by individual schools but that influence school performance, are identified. For example, student family background is the major environmental factor that impacts on school academic achievements (Ladd, 1996), and it also influences the allocation of Government grants in the New Zealand context[6]. To develop a reliable and valid performance measurement system, the relationship between environmental factor(s) and performance must be recognised and incorporated into the framework and models.

[6] The selection of inputs and outputs will be discussed in more details in section 2.3.3.

Figure 2.3 Steps to develop performance measurement framework for public schools



Adapted from Harrison et al. (2012)

Following the above steps, a refined framework with selected variables will be developed in section 2.3.3.

In summary, answers to the three questions initiated by Rouse and Putterill (2003, p. 795) are outlined in Table 2.2, which addresses the key issues for developing school performance measurement framework in the New Zealand educational context.

Table 2.2 Key issues for developing school performance measurement framework

Question 1 **Answer**	*Whose performance is observed?* The New Zealand public secondary schools.
Question 2 **Answer**	*What expectations are there for performance and how can these be transferred into measurable criteria?* Maximizing educational outcomes within budgeted school resources. Cost efficiency, academic efficiency and academic effectiveness of the education process are quantitative performance measures to satisfy the expectations of the primary stakeholder group (i.e., the Government and the MoE).
Question 3 **Answer**	*What attributes are to be measured and what are their relationships to the organizational environment?* School resources, academic inputs and outputs and educational outcomes are to be measured. They are all related to student family background, which is considered as the major environmental factor for the New Zealand education system (all the measures and the environmental factor will be further discussed in section 2.3.3).

2.3.3 Developing variables for school performance measurement framework

The input/output variables for an educational production model have been examined by a large number of studies in the education economic literature and there is no widely accepted view with regard to the specification of school performance measures or indicators (Hanushek & Raymond, 2005). However, as emphasized in section 2.3.2, performance measures should be linked to the objectives of primary stakeholders. That is to say, in New Zealand, school performance measures are expected to reflect the MoE’s main objective, being to “lift academic achievement and reduce inequality” (Ministry of Education, 2014c). The relevant priorities[7] are outlined as follows:

Priority 1. Strengthen literacy and numeracy teaching and learning;
Priority 2. Maximize the number of school-leavers with qualifications;
Priority 3. Support Mäori students’ success;
Priority 4. Support students’ engagement and success in education.
(Ministry of Education, 2011)

To refine the proposed performance measurement framework presented in Figure 2.2, this section will discuss the selection of variables for each efficiency element, in terms of what these variables measure and how they are linked to the MoE’s priorities. Moreover, the variable selection is also subject to the criteria of whether the variable is literature-based, or linked to the MoE’s priorities, or supported by available data.

2.3.3.1 Cost efficiency model

To measure school cost efficiency, the inputs that significantly relate to school performance are variables corresponding to school resources. School resources are ideally split into the resource of teachers and other operational resources to identify how much the efficiency of school resource utilization is related to teaching resources or other operational expenditures, given that both of them are consumed in the education process with the focus on student engagement and learning (Education Review Office, 2011). Therefore, they are all linked to Priority 4, and partially to Priority 1.

Performance indicators relating to resource of teachers could be reflected by the number of teachers (i.e., FTTE-full time teacher equivalent) or teachers’ salaries which reflect

[7] These priorities were extracted from the *Statement of intent 2011-2016*. A review of the MoE’s objectives since the introduction of NCEA (i.e., 2004-2011) suggested there had been no change in the nature of the MoE’s priorities during the period. As such, these priorities are used to select measures for school performance in any year after the introduction of NCEA.

the number of teachers, teachers' qualification levels and teaching experiences. Apart from teachers' salaries, other operational expenditures include learning resources and non-learning resources. Learning resources such as computers and sports equipment can be measured by the total cost of relevant expenditures (Hanushek, 1986; Rivkin, Hanushek, & Kain, 2005); non-learning resources consist of expenditures on administration staff salaries, school operation and property maintenance. In this study, teachers' salaries and other operational expenditures are included as the inputs for the cost efficiency model, given that the teacher numbers for the early years of the test periods (i.e., 2004-2005) are not available.

Since the Government and the MoE expect schools to maximize the number of students funded within the budgets, school roll is selected as the output of the cost efficiency model. Moreover, school roll is included as it is partially linked to Priority 2 and 4. In addition, school roll is in line with the MoE's goal of achieving the 'equality' of education, because school roll measures the number of students funded irrespective of the differences in student ethnic groups, gender and school deciles.

Of note, in New Zealand context, there has been little evidence supporting whether school size (measured by roll) can be fully controlled by school management. As a consequence of 'Tomorrow's Schools', school management are empowered more discretionary power to operate school but subject to a number of uncertainties relating to school zoning[8] and cost budgets. However, given that school roll is the unique and most reliable measure for the number of students funded, roll is selected as an output variable in the cost efficiency model.

2.3.3.2 Academic efficiency model

The proposed framework (Figure 2.2) demonstrates a network model that consists of three sub-processes namely cost efficiency, academic efficiency and academic effectiveness. The output(s) of the previous process is(are) exactly the input(s) of the next process in the network. Therefore, school roll as the output of the cost efficiency model is selected as the input of the academic efficiency model, which is used to measure how much academic achievement is obtained from students enrolled.

[8] As part of school enrolment scheme, school zones are set to prevent school overcrowding. That is to say "Only students living in the school's home zone are guaranteed a place at the school" (Ministry of Education, 2013).

With regard to the outputs of the academic efficiency model, a large number of studies advocate that examination results should be considered as the main indicator based on the fact that academic success is esteemed by the majority of the community, and a high level of examination results could satisfy the expectations of multiple stakeholder groups such as the Government and its agencies, parents, students and education providers (Hanushek, 1979). Even though the use of examination results is criticized for the reliability and validity of the examination instrument adopted, and also because the examination outcomes cannot reflect other school outputs such as social skills and sport achievements, the examination outcomes, especially the results of standardised achievement tests, are considered as the most reliable and comparable measures so long as the examination instrument is externally controlled and authenticated (Coe and Fitz-Gibbon, 1998). However, standardized tests are evaluated as having lower validity than curriculum-based examinations which are designed to achieve specific teaching goals (Hanushek, 1979).

In New Zealand, the standardized NCEA Level 1-3 examinations are taken by secondary school students from years 11 to 13. These examinations are centrally monitored and controlled by the NZQA which is independent from the MoE and education providers. Therefore, these examinations have a high level of reliability and comparability. However, the NCEA system allows schools to conduct internal assessment on students' performance through curriculum-based exams as per "unit standards" or "achievement standards", where internal assessment strategies by schools are potentially a source of variation. To improve the consistency and comparability of the school internal assessment system, since 2008, full-time moderators appointed by the NZQA take part in the moderation for up to 10% of internally-assessed students' work that is randomly selected. Furthermore, in August 2011, NZQA issued *Guide to Requirements for Consent to Assess for Schools*, together with the *Consent to Assess Against Standards on the Directory of Assessment Standards Rules 2011 (CAAS).* These documents are the guideline for schools to strength self-review processes to ensure an accurate and consistent internal assessment system. They are also the basis for the NZQA to evaluate the credibility of the school assessment system for the NCEA qualifications through Managing National Assessment (MNA) reports.[9] In addition, the

[9] A MNA report is issued to a school at least once every four years. Latest school MNA reports were firstly made available on line in November 2007 (New Zealand Qualifications Authority, 2015). However, the initial guidelines for the evaluation on the credibility of school NCEA assessment standards

NZQA can withdraw a school's right of internal assessment for a particular subject where necessary (New Zealand Qualifications Authority, 2016b). Thus, all these steps enable NCEA examinations to simultaneously have a high degree of reliability and validity (New Zealand Qualifications Authority, 2015).

Nevertheless, some researchers argue that using examination outcomes only to measure educational outputs is discordant with a comprehensive education system which should also cultivate social skills and psychological quality (Coe and Fitz-Gibbon, 1998). Still some researchers seek to resolve this issue by introducing additional school output measures such as school attendance rates or dropout rates, etc. (Hanushek, 1986; Ruggiero, 1996). Yet, this suggestion has been challenged because school attendance rates and dropout rates are process measures for obtaining educational outputs therefore they cannot represent outputs as such (Bosker and Scheerens, 1989).

Based on the above discussion, academic achievements are selected as the outputs of the academic efficiency model, being the number of NCEA Level 1, number of NCEA Level 2 and number of NCEA Level 3 passes. Given that students' achievements in literacy and numeracy are reflected by the requirements to achieve NCEA Level 1 qualification (New Zealand Qualifications Authority, 2015a), the MoE's Priority 1 is reflected by the number of NCEA Level 1 passes; moreover, Priority 2 is also reflected by the numbers of NCEA Level 1-3 passes which measure the number of qualifications achieved by students. The more the number of NCEA passes, the more the school leavers with qualifications. In addition, Priority 4 is indirectly associated with the numbers of NCEA Level 1-3 passes as well, because student engagement is related to academic efficiency, which measures a school's ability to get more students NCEA passes from the total students funded.

2.3.3.3 Academic effectiveness model

For the similar reason mentioned at the beginning of section 2.3.3.2, the outputs of the academic efficiency model are the exact inputs of the academic effectiveness model, being the number of NCEA Level 1, number of NCEA Level 2 and number of NCEA Level 3 passes.

As mentioned in section 2.3.2, in the proposed framework, the academic outputs measure school academic achievements (i.e., number of qualifications achieved) whilst

---

were developed in 2011. As such, MNA reports have little influence on the results of this study, given years after 2011 are not covered in the test periods.

the educational outcomes reflect the MoE's objectives (e.g., the number of UE for tertiary education); the amount of educational outputs may cause changes in the educational outcomes. As such, the academic effectiveness model measures how much educational outcomes are realised from the educational outputs.

In reference to Table 2.1, a number of outcome measures have been identified to reflect the MoE's strategic objectives that are presented in the MoE's statement of intent 2011. However, only the number of university entrance (UE) and the number of NCEA Level 3 Merit and Excellence (NCEA L3 M&E) endorsements are selected as school outcome measures because these two indicators represent high levels of academic performance, indicating that students are likely to succeed at first year university study or future employment (Scott, 2008). This is particularly in alignment with the Ministry's strategic goal, which is to achieve "Improved secondary-tertiary transitions between school, further education or work" (Ministry of Education, 2011, p.22). Moreover, these outcome measures are also related to Priority 2, given that NCEA L3 M&E were recorded for individual students since the beginning of the introduction of NCEA, and the certificates of NCEA Merit and Excellence endorsements for individuals were introduced in 2008 at qualification level (New Zealand Qualifications Authority, 2015). Thus, school leavers with NCEA L3 M&E certificates and UE qualification are more likely to gain success in tertiary study and future employment. In addition, Priority 4 is associated with the number of UE and the number of NCEA L3 M&E because with the evolution of the NCEA system, students are more encouraged than before to achieve UE qualification and Merit or Excellence endorsements, which also reflect students' success in New Zealand secondary education (New Zealand Qualifications Authority, 2015).

As a consequence of data unavailability or a possible distortion on school performance, other outcome measures in Table 2.1 are not considered as the outputs of the academic effectiveness model while the numbers of NCEA passes are included as the inputs. For example, the numbers of school leavers with NCEA Level 2 or above are not available. Moreover, the number of students staying on in education to age 17 may cause distortion on school performance, given that some students with poor academic performance stay on the education until age 17. Thus, this outcome indicator is not directly associated with the academic outputs reflected by the numbers of NCEA passes, and consequently may distort the scores of school academic effectiveness that measures the efficiency of how much academic outcomes are produced from the academic outputs (section 2.3.2).

In summary, to measure school effectiveness, the number of NCEA Level 1 passes, number of NCEA Level 2 passes and number of NCEA Level 3 passes are selected as the inputs whereas the number of UE and number of NCEA L3 M&E endorsements are included as the outputs.

2.3.3.4 Overall school efficiency

To evaluate the overall performance for each school, the overall school efficiency scores will also be produced by including teachers' salaries and other operation costs as the inputs with the number of UE and number of NCEA L3 M&E as the outputs. In this sense, cost efficiency, academic efficiency and academic effectiveness are actually the decomposition of the overall efficiency.

2.3.3.5 Environmental factor

The most important environmental factor which cannot be controlled by public school management is, student family background or socio-economic status (SES) (Ladd, 1996). This is a non-discretionary variable that has impact on all the efficiency models in the framework due to two reasons: first, it affects the level of government funding, which influences cost efficiency and overall efficiency; second, it affects the level of academic success variables, which impacts on academic efficiency, academic effectiveness and overall efficiency. As a result, cost efficiency, academic efficiency, academic effectiveness and overall efficiency are all subject to the adjustments for SES effect. This will be discussed further in section 2.4.2.

Student family background or SES could be reflected by parental education level and occupation, household crowding, welfare dependency or household income. Studies and NZQA statistics reveal that the family background of each individual student and that of his/her peers influence the level of his/her achievements (New Zealand Qualifications Authority, 2013; Feinstein and Symons, 1999; Haveman & Wolfe, 1995). Generally, schools catering for students from higher SES families achieve better academic outcomes compared to those with students from lower SES families. In New Zealand, family background is measured by a categorical decile[10] computed on the basis of census data. The MoE computes/updates the family background information for each individual student based on the census conducted every five years (Ministry of

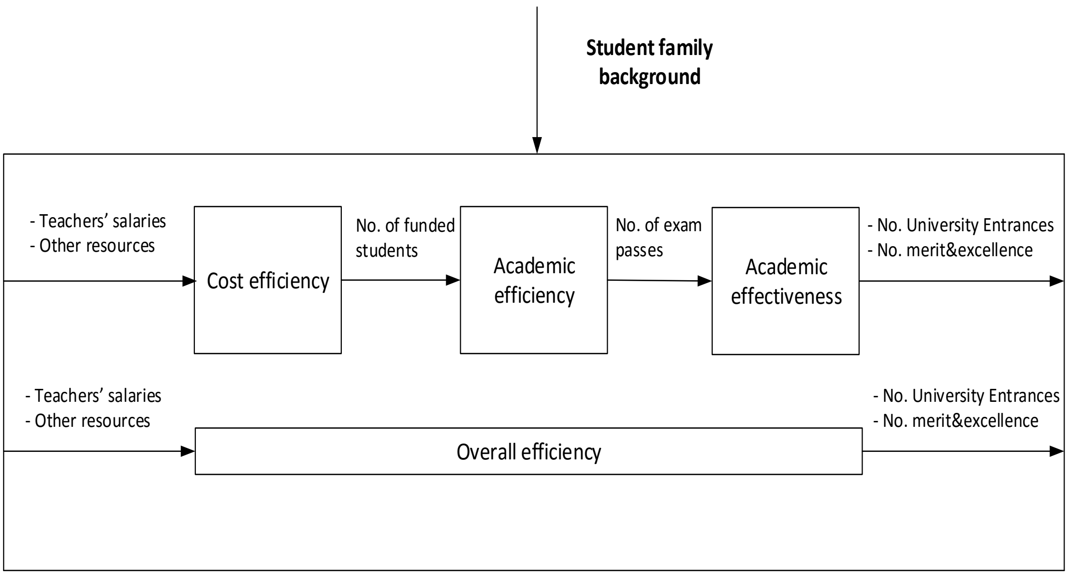
[10] Deciles are calculated as per five socio-economic indicators for a community: percentage of households with income in the lowest 20% nationally; percentage of employed parents in the lowest skill level occupational groups; household crowding; percentage of parents with no educational qualifications and percentage of parents receiving income support benefits. (Ministry of Education, 2016a)

Education, 2016a). Moreover, New Zealand school funding varies along with school decile, that is, schools with low decile tend to obtain higher levels of government funding while schools with high decile receive lower level of grants. The relationship between decile and school academic outputs, and that between decile and the Government grants per student will be further discussed in chapter 3.

Of note, it seems that there are no measures directly linked to Priority 3 regarding Mäori student support. However, as part of the population of New Zealand students, Mäori data have been included in each input or output and thus, all the measures selected in the framework are linked to Priority 3. The reason for not separating the Mäori group from other ethnic groups is that, at school level, the performance of the Mäori group is significantly associated with school Mäori density, which is considered another environmental factor. A further review suggested that school Mäori density is significantly associated with school SES decile, where this relationship cannot be one-to-one corresponded, given the categorical nature of school decile. This may technically cause distortion on the efficiency scores in the case that both decile and Mäori performance measures are considered in an efficiency model. Even so, the relationships between each efficiency and school Maori density will be examined in section 4.6 through a number of regressions.

Following the steps outlined in Figure 2.3, a refined framework has been developed (see Figure 2.4), which details the inputs and outputs for each efficiency model and specifies the environmental factor in the New Zealand education context.

Figure 2.4 School Performance Measurement Framework (Primary Stakeholders)



It is worth noting that with respect to the selection of performance measures, the refined framework not only incorporates the primary stakeholders' objectives but also is in line with the view that is popularly recognized in the economics of education literature: a conceptual model for school performance measurement should be designed by shaping the school objectives to best use school resources, while maximizing academic achievements and accomplishing educational outcomes to the greatest extent (Ruggiero, 1996).

In accordance with Figure 2.4, research question one (see section 1.2) as regards the components of a school performance model that measures both efficiency and effectiveness, could be answered so far. The model can be represented using a network model that consists of three inter-connected components, which successively measure cost efficiency, academic efficiency and academic effectiveness. The cost efficiency measures the efficiency of using school resources to support students funded; the academic efficiency measures the efficiency of student achievement in the NCEA Level 1-3 qualifications in all the students enrolled; the academic effectiveness measures the efficiency of converting the NCEA Level 1-3 passes to the UE passes and NCEA L3 M&E endorsements. All the above components are subject to the adjustments for the effect of student family background.

In summary, the framework depicted in Figure 2.4 is supported by the primary stakeholder theory and the Ramanathan's Framework. The rationale of this framework is to identify what facet(s) of performance are to be measured (e.g., efficiency and effectiveness) as per the expectation of primary stakeholders(s), what measures are directly controlled by the organization and how to link model measures to the objectives of primary stakeholder(s). This ensures the applicability of this framework to the New Zealand secondary education system, which is relatively homogeneous in the sense that the majority of schools are state-funded with many obligations imposed by the MoE. Given the MoE functions as the sole agency of the Government that is the primary stakeholder in New Zealand education sector, the inputs and outputs for each efficiency model in the framework are selected in alignment with the MoE's objectives.

Section 2.4 reviews the literature on the quantitative research techniques used for school performance measurement.

2.4 Quantitative approaches to school performance measurement

According to the final framework shown in Figure 2.4, there are four efficiencies (i.e., cost efficiency, academic efficiency, academic effectiveness and overall efficiency) to be produced for school performance measurement. Given that they have the same nature in terms of calculating efficiency scores from inputs to outputs with the same environmental factor(s), sections 2.4.1 and 2.42 search for a quantitative method that is common to modelling any of the above efficiencies in the New Zealand educational context. Section 2.4.3 discusses the formulation of Malmquist index that is calculated to measure school productivity changes. Section 2.4.4 explains how to implement the final framework based on the selected approach.

2.4.1 Overview

Given there is scarce guidance on the ideal empirical approach to measure school performance (Krueger, 1999), the quantitative methods in most common usage are regression analysis models and mathematical optimisation models. Examples of regression models are ordinary least squares (OLS), corrected ordinary least squares (COLS) and stochastic frontier analysis (SFA) (Hanushek, 1979; Krueger, 1999), while the representative of mathematical optimisation models is Data Envelopment Analysis (DEA) (Ruggiero, 1999).

The preconditions of applying regression analysis such as the requirement of linearity and independence between variables may not be satisfied in a school performance measurement system, and this may restrict the applicability of the regression models with regard to measuring school performance. For example, the large variety of school expenditure levels cannot guarantee linearity; moreover, the regression approach estimates relative efficiency values by using a statistical average of performance as a benchmark rather than producing a frontier of “best performance”. Therefore, this method is less suited to provide targets or efficient peers[11] for inefficient schools to improve performance. However, it may be argued that SFA can generate a frontier so called “deterministic frontier” on which all the efficiency values are produced by removing the effect of statistical noise and inefficiency, the efficiency figures are calculated relative to the figures on the frontier for each school. However, SFA requires the specification of the mathematical function for the production process, while this is

[11] For a school under observation, its peers refer to the schools with similar characteristics (SES decile and size etc.), but better efficiency.

hard to be fulfilled due to the lack of knowledge with respect to how educational inputs relate to outputs (Harrison and Rouse, 2014).

In contrast to these regression models, DEA produces a frontier of “best performance” for a collection of organizations that have similar attributes and strategic goals. It is therefore applicable in New Zealand, given the homogenous education system and the need to develop a benchmark to distinguish under-performing schools. For the regression models, DEA generates relative efficiency but uses the frontier as the baseline so as to supply a composite efficiency figure for each school. Also, DEA provides efficient peers and projected targets for each inefficient school together with the relative weight of each peer based on how similar the environment that the peer operates in comparison with the school. This allows a comparison of performance for all the schools being evaluated, and also suggests the peers and targets for each inefficient school to improve its performance. Moreover, unlike the SFA approach, DEA does not require the pre-specification of a production function or the price/weight data for any inputs or outputs in a performance measurement model, and this makes DEA more preferable to SFA (Harrison and Rouse, 2014).

Nevertheless, compared with other models (e.g., regression models), DEA models are more sensitive to measurement error and improper selection of model variables. Further, its performance is unstable in the case that the samples vary significantly in terms of data attributes, even though they are selected from the same population. Thus preferably, DEA should be used for production units with similar attributes in relatively big samples to overcome endogeneity issues (Ruggiero, 2003); sample size should be at least twice of the number of inputs and outputs combined (Cook, Tone and Zhu, 2014). In addition, the distortion caused by measurement errors on DEA results decreases significantly where aggregated data is applied. As such, accumulated data at school level is used as opposed to individual student level (Ruggiero, 2006). However, the accumulation might cover different types of inputs or outputs and this might distort DEA results to a certain extent (Barnum and Gleason, 2006). Therefore, the accumulation across inputs/outputs must be controlled to a tolerable level through a proper design of variable types for the DEA model.

Apart from the endogeneity of variables that can be overcome by using a large sample, heterogeneity issues may occur where discrepancy in individual features of schools exists, typically, student family background. However, the heterogeneity problem relating to production process could be minimized in the New Zealand context due to the high level of standardization of the education process in terms of curriculum, teachers' qualification, reward system, examination standards, central administration and financial reporting (Alexander et.al, 2010). According to Harrison (2008), the heterogeneity issues related to student family background could be resolved by the proposed DEA model which employs family background variables to control these discrepancies. This will be discussed further in section 2.4.2.

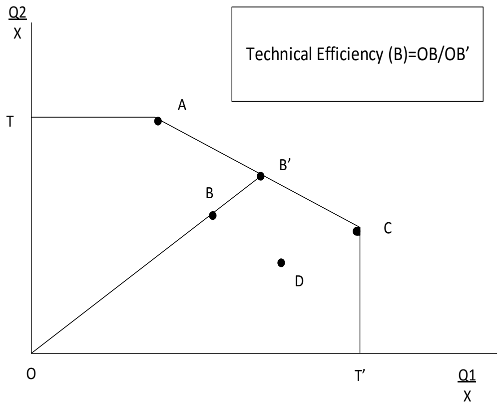
In summary, there is little guidance about which empirical approach is most appropriate to measure school performance. The most common methods are regression analysis models represented by SFA and mathematical optimisation models represented by DEA. The major reason for choosing DEA is that, this model does not require any pre-specification for school production function and can work well over large samples even in presence of endogeneity of variables. Other statistical issues such as heterogeneity and measurement errors can be overcome by a proper design of DEA model in terms of variable types and control variables in the model.

2.4.2 DEA fundamentals

DEA was introduced by Charnes, Cooper and Rhodes (1978) based on the earlier work in relation to the theoretical fundamentals for productive efficiency measurement (Farrell 1957; Debreu, 1951; Koopmans, 1951). Farrell (1957) defines the concept of "technical efficiency" which refers to a firm's success in maximising the production of outputs under a given level of inputs. The technical efficiency of a production unit, so called "Decision Making Unit" (DMU), is calculated through the increase in proportional outputs to inputs (output-orientation) or decrease in proportional inputs to outputs (input-orientation), that is needed for an inefficient DMU to be technically efficient. That is, the distance from the DMU to the production frontier, which reflects the best performance for all DMUs that have 100% technical efficiency.

Using an example of four DMUs (A, B, C and D), Figure 2.5 demonstrates how the technical efficiency for DMU B is calculated, assuming each of these DMUs uses one input (X) to produce two outputs (Q_1, Q_2). The efficiency measurement is output-oriented (getting the most outputs from the given inputs). The production frontier is represented by the isoquant TT'. A and C are on the production frontier therefore they are all 100% technically efficient. Technical efficiency of DMU B is measured by OB/OB' and BB' reflects its degree of technical inefficiency.

Figure 2.5. DEA: measures of efficiency (output-orientation)



DEA identifies the production frontier through a Linear Programming Model (LPM). Equation 2.2 displays a typical DEA model for both input orientation and output orientation under both constant returns to scale (CRS) and variable returns to scale (VRS). These terms are explained below.

Equation 2.2 DEA Envelopment Form

	Input orientation	Output orientation
Constant Return to Scale (CRS)	Min $h_o = \theta$ s.t. $\sum_{j=1}^{n} x_{ij}\lambda_j \leq \theta x_{io}$ $i=1,\ldots,m$ $\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{ro}$ $r=1,\ldots,s$ $\lambda_j \geq 0,$ $j=1,\ldots,n$ θ_o unrestricted	Max $h_o = \theta$ s.t. $\sum_{j=1}^{n} y_{rj}\lambda_j \geq \theta y_{ro}$ $r=1,\ldots,s$ $\sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{io}$ $i=1,\ldots,m$ $\lambda_j \geq 0,$ $j=1,\ldots,n$ θ_o unrestricted
Variable Returns to Scale (VRS)	Min $h_o = \theta$ s.t. $\sum_{j=1}^{n} x_{ij}\lambda_j \leq \theta x_{io}$ $i=1,\ldots,m$ $\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{ro}$ $r=1,\ldots,s$ $\sum_{j=1}^{n} \lambda_j = 1$ $\lambda_j \geq 0,$ $j=1,\ldots,n$ θ_o unrestricted	Max $h_o = \theta$ s.t. $\sum_{j=1}^{n} y_{rj}\lambda_j \geq \theta y_{ro}$ $r=1,\ldots,s$ $\sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{io}$ $i=1,\ldots,m$ $\sum_{j=1}^{n} \lambda_j = 1$ $\lambda_j \geq 0,$ $j=1,\ldots,n$ θ_o unrestricted

Where, $\mathbf{y_{ro}}$ and $\mathbf{x_{io}}$ are vectors of outputs and inputs for the DMU under observation, which is an element of the matrices of n × s outputs (y) and n × m inputs (x) for n DMUs. λ represents the scalar applied to DMU_0's peers while θ denotes the scalar measure of efficiency.

As mentioned above, DEA models are categorized as input orientation and output orientation models. An output orientation model calculates the efficiency score of output production under a constant level of input resources. It is particularly suitable for public schools which are operating within the budget approved by the Government. In contrast, an input orientation model computes the efficiency scores for the use of input resources under a constant level of outputs.

DEA generates an efficiency frontier using the efficient DMUs (100% technical efficiency) from the sample. All the combinations of inputs, outputs and efficiencies of the DMUs in the sample are enveloped by the frontier (Charnes et. al, 1978). Where the productive efficiency is dependent on returns to scale, the production frontier can embrace either constant returns to scale (CRS)[12] or variable returns to scale (VRS).[13] Compared to CRS, VRS incorporates different scale/size of each individual DMU into a performance evaluation process to produce efficiency scores within the sample.

As mentioned in section 2.3.3, there has been little evidence suggesting whether school management can fully control school size due to the restriction of school zoning and cost budgets. Therefore, both CRS and VRS efficiency scores will be calculated and compared in this study. The concept of scale efficiency is introduced to reflect the proportion of the DMUs which operate at the most productive scale in the sample. Scale efficiency is computed as the ratio of CRS efficiency to VRS efficiency.

As mentioned in section 2.3.4, the difference in student family backgrounds may result in heterogeneity that has a negative impact on DEA efficiency scores. The variable of student family backgrounds measured by SES decile in New Zealand, is considered as a non-discretionary variable for school management due to the restriction of school-zoning. Moreover, research has revealed that students' academic performance is significantly related to the SES decile. For example, higher academic achievement relates to students with prosperous families (Jensen and Seltzer, 2000). Accordingly, a categorical DEA model is selected to handle non-discretionary variables by grouping the DMUs into different SES decile categories where the decile variable is categorical/ordinal; each DMU is evaluated against its group-mate DMUs together with those in the categories under more difficult conditions (Harrison and Rouse, 2014).

[12] CRS refers to the circumstance that output increases by the same proportion as that of the increase in all levels of inputs. CRS can be achieved by optimizing a DMU's operation scale.

[13] A VRS efficiency frontier consists of three parts; increasing returns to scale (IRS), when the amount of output increases by more than the proportional increase in inputs; CRS and decreasing returns to scale (DRS), when the amount of output increases by less than the proportional increase in inputs.

The categorical DEA approach ensures that a DMU is not directly evaluated against or compared with those in higher categories (under more favourable conditions). Meanwhile, the peer group with full efficiency for each inefficient DMU is identified. In New Zealand, the MoE publishes SES decile for each school periodically, which is calculated by incorporating a range of measures regarding student family backgrounds in connection with student achievements. The decile is ordinal data ranged between one and ten, where one indicates the highest proportion and ten represents the lowest proportion, of students in the school are from the most deprived areas. As such, the categorical nature of the decile in New Zealand enables the application of the categorical DEA model (Harrison and Rouse, 2014).

A standard categorical model (VRS, output orientation) is presented in Equation 2.3.

Equation 2.3 Categorical DEA model (VRS, output-orientation)

$$\text{Max } h_0 = \theta$$

$$\text{s.t.} \sum_{j=1}^{n} \lambda_j Y_{rj} \geq \theta Y_{ro} \quad r=1,\ldots,s$$

$$\sum_{j=1}^{n} \lambda_j X_{ij} \leq X_{io} \quad i=1,\ldots,m$$

$$\sum_{j=1}^{n} \lambda_j d^{C_k}{}_{ij} \leq d^{C_k}{}_{io} \quad k=1,\ldots,t;\ C_k=1,\ldots,C\text{-}1$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

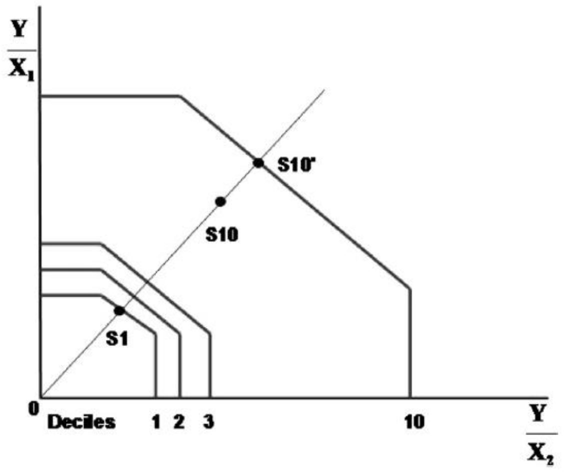
$$\lambda_j, X_{ij}, Y_{rj} \geq 0;\ j=1,\ldots,n$$

Adapted from Banker & Morey (1986) and Harrison, Rouse & Armstrong (2012)

Note, in the third set of constraints, $d^{C_k}{}_{io}$ is a dummy variable which represents categories relating to environmental difficulty, where the superscript C signifies the category level (e.g. school decile 1-10) and k denotes the k^{th} category variable (where more than one category variables exist, i.e., t>1).

Figure 2.6 shows how a categorical DEA model is applied to the New Zealand public schools based on the decile categories, assuming there are two inputs (X_1,X_2) that produce one output Y. The production frontier for each decile category is computed based on the schools in the same decile together with those in the lower decile apart from decile one, where the efficiency score is determined by the schools only in decile one. For example, S1 is a school in decile one with 100% efficiency (sitting on the frontier), and the decile one frontier is determined only by the decile one schools. In contrast, S10 is a decile ten school, and its efficiency is measured against the decile ten frontier. It is not 100% efficient because its efficiency is lower than the decile ten frontier, which is computed based on all the schools from decile one to decile ten. Albeit S10 produces more outputs than S1 under the same inputs level, however, S10 is less efficient than S1 against its own decile frontier (S10*) (Harrison and Rouse, 2014).

Figure 2.6 Applying a categorical DEA model



Source: Harrison and Rouse (2014, p.13)

2.4.3 Malmquist index

Since research question two aims to address productivity changes in New Zealand school performance, the Malmquist technique is employed in this study to measure productivity changes in school performance efficiencies. As a derivative of DEA technique, Malmquist index is normally calculated to examine the impact of policy changes over time on the efficiency of DMUs within a sample. It is obtained from ratios of distance functions, which were developed by Caves, Christensen and Diewert (1982). Since an output-orientation DEA model is preferred to measure public school performance in this study (details are discussed in section 2.4.4), only the output distance function is displayed as follow:

Equation 2.4 Output distance function

$$D_0^t(y^t, x^t) = \min\{\boldsymbol{\delta}: (x^t, \frac{y^t}{\delta}) \in S^t\}$$

Where, for time period *t*, S^t consists of all the combinations of input (x) and output (y) being feasible in the production process. Actually, the output distance function reflects the maximum proportional radial expansion of outputs for a particular set of inputs (Caves et.al, 1982). In practice, $D_0^t(y^t, x^t)$ is reflected by the efficiency at time *t*.

According to Färe, Grosskopf, Lindgren and Roos (1989, 1994), the Malmquist index can be decomposed into two parts: change in efficiency (i.e., movements of a DMU's relative efficiency) and change in technology (i.e., frontier-shift). Assuming the observation is for the change from time *s* to *t*, the Malmquist index can be decomposed as follows:

Equation 2.5 Malmquist index (output orientation)

Malmquist index=Efficiency Change× Frontier-Shift

Where, Efficiency change=$\frac{D_0^t(x^t,y^t)}{D_0^s(x^s,y^s)}$; Frontier-shift=$[\frac{D_0^s(x^t,y^t)}{D_0^t(x^t,y^t)}\frac{D_0^s(x^s,y^s)}{D_0^t(x^s,y^s)}]^{1/2}$

In the frontier-shift formula, $D_0^s(x^t, y^t)$ measures the efficiency at time *t* against the frontier at time *s*; similarly, $D_0^t(x^s, y^s)$ measures the efficiency at time *s* against the frontier at time *t*. Both $D_0^s(x^t, y^t)$ and $D_0^t(x^s, y^s)$ are also called cross-efficiencies.

An index number over one indicates an increase, a number equal to one means no change and a number below one represents a decrease in efficiency.

The Malmquist index can be produced using CRS or VRS mode. However, research suggests only CRS works well for the calculation of Malmquist indices (Färe et.al, 1989, 1994), because VRS mode may cause bias in the calculation of productivity change when dealing with different types of returns to scale (Grifell-Tatjé and Lovell, 1995). As such, Malmquist index and the two parts of its decomposition will be calculated using CRS model.

2.4.4 Using DEA to measure New Zealand school efficiencies

The DEA approach has been selected to model cost efficiency, academic efficiency, academic effectiveness and overall efficiency that reflects the different facets of public school performance in New Zealand, given that each DEA model in the framework produces a composite efficiency score for each school by incorporating multiple inputs and outputs without the necessity to formulate a production function. In the New Zealand context, an output oriented DEA model is applied for the calculation of all the above efficiencies, given that the MoE expects each public school to maximize academic outputs and outcomes from the budgeted resources. Scale effect is considered by calculating efficiency scores under both CRS and VRS modes due to the absence of knowledge in the literature on the relationship between school operation scale and school performance.

In regard to the variable settings in the DEA models, inputs and outputs are measured by absolute values rather than averages or ratios to avoid possible distortions on DEA results (e.g., number of NCEA passes rather than pass rate) (Dyson et al., 2000; Harrison, 2008). With respect to the non-discretionary factor being school SES decile which reflects student family background characteristics, the categorical DEA approach is selected to generate school efficiency scores for each school against the production frontier for the same decile, where the frontier is identified based on all the schools in the same or less favourite socio-economic conditions.

In summary, to model any of the efficiency elements demonstrated in Figure 2.4, an output- orientated DEA categorical model using decile categories will be run under both CRS and VRS modes.

Figure 2.4 shows that in the proposed framework, the sub-processes in sequence measured by cost efficiency, academic efficiency and academic effectiveness comprise the whole educational process for New Zealand public schools, and also demonstrates the relationships between them. That is, the outputs of the cost efficiency model are exactly the inputs of the academic efficiency model, where the outputs of the academic efficiency model just contribute to the inputs of the academic effectiveness model. This is a typical structure of a network DEA model, which uses the outputs of one DEA model as the inputs for the next one in the network; the transformation of original inputs into final outputs is decomposed into several transitional stages with the production of intermediate outputs (Färe and Grosskopf, 2000).

Unlike Harrison (2008) and Alexander et. al (2010) which measure school performance using a single DEA model, this study employs a network DEA model to capture school performance information from multiple dimensions by including cost efficiency, academic efficiency and academic effectiveness models into the performance measurement process. In doing this, the sources of inefficiency can be identified in parts of the process, and the interests of different stakeholders can be reflected by the efficiencies produced from these inter-connected and complementary sub-models (Amado, Santos and Marques, 2012). For example, with respect to school evaluation, the MoE is concerned with all the above efficiencies while students and parents are more interested in academic efficiency and academic effectiveness.

To examine the difference between the network DEA and the single DEA with respect to the efficient schools identified, school overall efficiency is also modelled by incorporating the original inputs and the final outputs of the process. The comparison of results will be discussed in chapter 4.

2.5 Summary

Public school performance is expected to be reflected by the efficiency of resource utilization and the effectiveness of objective achievement in the organisation, and the performance measurement framework should incorporate the objectives of primary stakeholder group, being the Government and the MoE in the New Zealand educational context. Thus, underpinned by Ramanathan's framework, a performance measurement framework for New Zealand public schools is developed (Figure 2.4), which decomposes the whole educational process into three processes measured by cost efficiency, academic efficiency and academic effectiveness, where all the input/output variables in the framework are under the direct control of school management and they are all linked to the MoE's objectives.

Compared with other popular quantitative methods for performance measurement, DEA is considered more appropriate to measure the efficiencies displayed in the framework, given that DEA combines multiple input and output measures to produce composite efficiency scores for each school without the need to define the production function and these scores can be used to compare and rank school performance.

According to the specific features of the New Zealand education system, an output-oriented categorical DEA model will be used to calculate the efficiencies included in the framework. The school SES decile is incorporated into the categorical model to adjust for the impact of student socio-economic status on efficiency scores.

To implement the framework, a network DEA model is employed to measure school performance via a series of inter-connected sub-models, arranged in the order of cost efficiency, academic efficiency and academic effectiveness. The variables of each sub-model are selected in alignment with the MoE's priorities.

To serve research question two, the Malmquist index and its decomposed parts will be used to measure the productivity changes in school performance efficiencies.

Chapter 3 discusses the research methodology, data collection and sample formation for this study.

Chapter 3 Research Methodology and Data Collection

3.1 Introduction

A positivist approach is adopted in this research based on the assumption that “meaning exists in objects independently of any consciousness” (Crotty 1998, p.10). That is to say, knowledge is unaffiliated to the researcher’s beliefs; “realism” and “objectivism” are the key characteristics of a positivist approach. The positivist methodology is valid as the research purpose is to examine New Zealand school performance based on the belief that school performance can be measured objectively, and the research questions are examined by DEA (for school performance measurement) and its derivative (i.e., Malmquist approach) that are completely in harmony with the positivist research paradigm.

This chapter concerns the research design and method for constructing the network DEA model to measure New Zealand school performance. Following the introduction, section 3.2 discusses the definitions for all the variables in the network DEA model. Section 3.3 describes the procedures for data collection and two test periods are determined in this section (see subsection 3.3.4), being 2004-2006 and 2009-2011. Section 3.4 addresses sampling issues and two research samples are formed for comparison, being the full sample which contains both expanded schools (years 7-15) and pure secondary schools (years 9-15), and the pure secondary school sample (years 9-15).

3.2 Definitions for DEA model of school performance measurement

Following the discussion in chapter 2, the variables for the proposed network DEA model are detailed in Table 3.1. The variables for each efficiency model (i.e. cost efficiency, academic efficiency and academic effectiveness, overall efficiency), are described in section 3.2.1.

Table 3.1 Variables of the proposed DEA models for school performance

Inputs	Non-discretionary variables	Outputs
Cost efficiency model 1. Teachers' salaries; 2. Operational expenditures excluding teachers' salaries	Social-economic (SES) decile	School roll
Academic efficiency model School roll	Social-economic (SES) decile	1. Number of NCEA Level 1 passes 2. Number of NCEA Level 2 passes 3. Number of NCEA Level 3 passes
Academic effectiveness model 1.Number of NCEA Level 1 passes 2.Number of NCEA Level 2 passes 3.Number of NCEA Level 3 passes	Social-economic (SES) decile	1. Number of UE; 2. Number of NCEA L3 M&E
Overall efficiency model 1. Teachers' salaries; 2. Operational expenditures excluding teachers' salaries	Social-economic (SES) decile	1. Number of UE; 2. Number of NCEA L3 M&E

3.2.1 Inputs and outputs in the network DEA model

3.2.1.1 School resources

School resources being utilized to produce school outputs/outcomes are considered as the primary inputs of cost efficiency and overall efficiency models. For the purpose of analyzing how different inputs impact on the production of outputs, two school resources are included as the inputs: teachers' salaries and operational expenditures excluding teachers' salaries. All these school resources are under the direct control of school management.

Operational expenditures consist of expenditures on learning resources and expenditures on other resources. Expenditures on learning resources relate to curricula, teachers' salaries, equipment repairs, ICT leases, staff development, attached teacher costs, library resources and other learning resources. Expenditures on other resources include administration expenses and property expenses. The operational expenditures split into teachers' salaries and other operational expenditures are selected as the inputs for cost efficiency and overall efficiency models based on the principle that they are all directly related to the teaching and learning process. Specifically, administration expenses relate to school operation expenses such as BoT expenses, consumables, administrative staff salaries, ACC and consultancy fees etc.; property expenses include all the costs relevant to maintaining the physical environment for student learning and teaching such as caretaking and cleaning consumables, consultancy and contract servicers on property, heat, light and water etc.

School expenditure categories excluded from the inputs for cost efficiency and overall efficiency models are:

- Depreciation and amortization
  These amounts represent the capital/intangible assets consumed during the education process. However, their reasonableness and accuracy are in question due to the fact that there is a large variety of school buildings and equipment in terms of the initial values and ages in the sample. Therefore, the depreciation or amortization figures cannot represent a fair and reasonable proxy for school annual capital consumptions.
- Hostel expenses and international students' expenses
  These expenses relate to the activities of attracting more international students to generate income for school, therefore they are not directly relevant to education production.
- Loss on asset disposal
  This term relates to school financial profit, rather than educational outputs.
- Impairment
  This expense is directly associated with the decrease in the book value of school capital assets, compared with the market value rather than the annual capital consumption in education process. It is excluded for similar reasons as depreciation.
- Finance costs and local funds expenses
  These terms reflect the expenses of fundraising and financing, therefore, they are directly connected with the increase in school capital value rather than educational activities.

It is worth mentioning that teacher number (number of teacher full time equivalent, known as FTTE) is not selected to reflect teaching resource due to the unavailability of the data for 2004 and 2005 in the test period. Alternatively, teachers' salaries in dollar is included as a DEA input to measure teaching resources.

In New Zealand, there are three types of schools providing secondary education: composite schools (years 1-15), expanded secondary schools (years 7-15) and pure secondary schools (years 9-15 or 11-15). This research excludes composite schools due to the concern for the homogeneity of each sample, given the educational process between primary and secondary schools is significantly different in terms of government funding rate and curriculum settings. However, as there are still two types included in the sample, it is necessary to adjust expenditures to remove the effect of non-secondary education provision.

The allocation rate of government grants for school operation varies with the education level of students (i.e. intermediate or secondary level) provided by schools. However, expenditures of expanded secondary schools provided by the MoE are not separately presented in correspondence to different education levels; therefore, it is essential to split the expenditures into two parts for each expanded school: amount for secondary education and amount for intermediate education. Only those resources related to secondary level are included into the network DEA model (Ministry of Education, 2014k).

The steps[14] for the expenditure splitting are listed below:

For each year observed (i.e., each year in 2004-2006 and 2009-2011),

- In Step One, calculate teachers' salaries per student (teacher$/student) and other operational expenditures per student (other$/student) for all the pure intermediate schools (years 7-8). Because the average amount of expenditures varies with school roll, these measures are calculated for school roll under 500 and over 500. Compute average intermediate teacher$/student respectively using 5% trimmed mean of teacher$/student for all the pure intermediate schools, then do the same to obtain average intermediate other$/student.
- In Step Two, for each expanded school in the sample, collect roll data of years 7-8 (roll 7-8). Calculate teachers' salaries for intermediate education by multiplying average teacher$/student obtained from Step One and roll 7-8, then do the same to get other operational expenditures for intermediate education.
- In Step Three, for each expanded school in the sample, deduct the teachers' salaries for intermediate education from the total teachers' salaries; similarly, subtract the

[14] Following Harrison (2008), a regression was also run to obtain the ratio for the expenditure splitting between intermediate and secondary education. However, the ratio produced from the regression was inconsistent with the MoE's funding formula, showing that intermediate students (years 7-8) consumed more expenditure than secondary students (years 9-15).

amount of other operational expenditures for intermediate education from the total other operational expenditures.

The summary of average expenditures per intermediate student for each year observed is listed in Appendix II. To verify whether the splitting approach is valid, two *t*-tests were undertaken to compare the mean value of total expenditures per secondary student between expanded schools and pure secondary schools in the sample after removing the expenditures for intermediate education. Results for the t-tests are recorded in Table 3.2, showing there are no significant differences in mean expenditures per secondary student for both 2009-2011 and 2004-2006 periods. This suggests that after the splitting, expenditures per secondary student for the expanded schools is at the same level as that for the pure secondary schools with similar roll size, consistent with the expectation that in New Zealand the allocation of Government grants for schools is highly standardized. In this sense, Table 3.2 provides evidence to support the validity of the expenditure splitting approach for expanded schools.

After the splitting, the ratio of total expenditures between intermediate and secondary education is 21:79 for expanded schools in 2004-2006, and 22:78 in 2009-2011.

Table 3.2 Comparison of mean cost/secondary student between expanded and pure secondary schools after expenditure splitting

	N	Cost/secondary student		*t*-value	Degree of freedom	Significance (2-tailed)
2009-2011		**Mean**	**Std. Dev.**			
Expanded	226	9222.99	2026.69	1.44	362.89	0.151
Secondary	375	9001.15	1441.86			
2004-2006						
Expanded	222	7142.59	1438.25	1.50	394.59	0.135
Secondary	294	6970.10	1077.61			

Note:

1. Results were produced without equal variances assumed;
2. The *t*-test for 2009-2011 period was based on schools with roll under 1,000, as such 95% of the expanded schools were included in the sample; the test for 2004-2006 period focused on schools with roll under 833, therefore 97% of the expanded schools were contained within the sample. Sampling for these *t*-tests was based on the rule to ensure the expanded and pure secondary schools in the sample were with similar size and balanced distribution, given that cost per student is significantly associated with school roll;
3. All the cost amounts in this table have been indexed to adjust inflation effect.

In parallel with the expenditure splitting for secondary education, the roll variable is also expected to solely reflect the number of years 9-15 students. Further, to reduce any possible distortion on data analysis caused by inflation during the test periods, all the financial inputs for the DEA model are indexed by applying the Gross National Expenditure (GNE) deflator, given that the GNE price deflator reflects the expenditures consumed by both households and government entities, and also reflects gross capital formation (Trading Economics, 2014).

Since the school financial year ends on 31 December, the average GNE deflator for each year during the test periods is used to index yearly financial inputs. The base year is 2004 (Statistics New Zealand, 2015).

3.2.1.2 Number of NCEA exam passes

As mentioned in chapter 2, the numbers of NCEA Level 1-3 passes are selected to represent school academic outputs. Given that these numbers cannot be collected directly from the NZQA statistics, these data are computed based on the data from the files known as 'Qualification-Statistics-School-yyyy' [15] published on the NZQA's official website.

As described in section 2.2.2.2, NCEA Level 1- Level 3 qualifications are credit-based and can be taken in any of years 11-13 (New Zealand Qualifications Authority, 2016). Therefore, the total number of NCEA passes for each NCEA level (i.e., Level 1-3) is the sum of NCEA passes for this level from all the student year levels (i.e., years 11-13). For any NCEA exam year, the current year achievement rates on NCEA Level 1-3 are split by student year levels (i.e., years 11-13) and published on the NZQA website. However, to calculate the number of NCEA Level 1-3 passes for each year level, school roll data for each year level are required. They were collected from the MoE's website.

For each school under observation, the numbers of NCEA Level 1-3 passes are calculated as below:

Number of NCEA Level 1 passes= Current Year Achievement Rate Roll Level 1 year11×roll year11+ Current Year Achievement Rate Roll Level 1 year12×roll year12+ Current Year Achievement Rate Roll Level 1 year13×roll year13;

[15] 'yyyy' denotes any year in the two test periods (i.e., 2004-2006 and 2009-2011). For example, for year 2004, the corresponding data file's name is 'Qualification-Statistics-School-2004'

Number of NCEA Level 2 passes= Current Year Achievement Rate Roll Level 2 year11×roll year11+ Current Year Achievement Rate Roll Level 2 year12×roll year12+ Current Year Achievement Rate Roll Level 2 year13×roll year13;

Number of NCEA Level 3 passes= Current Year Achievement Rate Roll Level 3 year11×roll year11+ Current Year Achievement Rate Roll Level 3 year12×roll year12+ Current Year Achievement Rate Roll Level 3 year13×roll year13.

3.2.1.3 Number of university entrance qualifications and number of NCEA Level 3 Merit and Excellence endorsements

As discussed in chapter 2, the number of UE together with the number of NCEA Level 3 M&E are selected as the outputs for the academic effectiveness model. Because the threshold for NCEA Level 3 qualification is used for the UE achievement (New Zealand Qualifications Authority, 2013), UE is mostly obtained by year 13 students, and only a small number of UE passes are achieved by year 12 or 11 students. The calculation of UE passes is similar to that for NCEA Level 3 passes. The number of UE is computed based on the information provided by the file namely ‘Qualification-Statistics-School-yyyy’.

The NCEA endorsements (i.e., Pass, Merit and Excellence) are also credit-based. For each school, the number of NCEA L3 M&E is the sum of the number of NCEA Level 3 Merit and number of NCEA Level 3 Excellence. However, these figures cannot be acquired from the NZQA website, as only the endorsement achievement rates for NCEA Level 1-3 are provided, split by student year levels. To obtain the number of NCEA L3 M&E, all of the number of NCEA Level 3 passes from each student year level are required, and the number of NCEA L3 M&E equals the achievement rate for this endorsement multiplied by the number of Level 3 passes; the total number of NCEA L3 M&E is the sum of that from all the year levels.

Formulas for the calculation of UE passes and the number of NCEA L3 M&E are listed below:

Number of UE= Current Year Achievement Rate Roll UE year11×roll year11+ Current Year Achievement Rate Roll UE year12×roll year12+ Current Year Achievement Rate Roll UE year13×roll year13;

Number of NCEA Level 3 Merit= Current Year Achievement Rate Level 3 Merit year11× number of year11 Level 3 qualification passes+ Current Year Achievement Rate Level 3 Merit year12× number of year12 Level 3 qualification passes+ Current Year Achievement Rate Level 3 Merit year13× number of year13 Level 3 qualification passes[16];

Number of NCEA Level 3 Excellence= Current Year Achievement Rate Level 3 Excellence year11× number of year11 Level 3 qualification passes+ Current Year Achievement Rate Level 3 Excellence year12× number of year12 Level 3 qualification passes+ Current Year Achievement Rate Level 3 Excellence year13× number of year13 Level 3 qualification passes;

Number of NCEA L3 M&E= Number of NCEA Level 3 Merit+ Number of NCEA Level 3 Excellence.

3.2.2 Non-discretionary variables

As mentioned in chapter 2, the most important non-discretionary variable is the student background characteristic, which is measured by social-economic (SES) decile. The decile incorporates a variety of information in relation to equivalent household income, parents' occupation, household crowding, parents' educational qualifications, income support payments received by parents (Ministry of Education, 2014m). A school decile reflects the extent to which the school enrolls students from low level socio-economic communities, with a decile one school holding the highest percentage of deprivation and a decile ten school having the lowest. A decile ranking is allocated to each state-owned, state-integrated and special school and all the decile data are reviewed every five years. Of note, deciles were developed to "assistant in differential funding to schools" (Hattie, 2002).

A number of research findings reveal that school academic performance is positively related to school SES decile (Harrison, 2008; Hattie, 2002). Moreover, statistics from NZQA website also show the positive relationship. To examine the actual relationship between school academic achievement and decile, mean and standard deviation of

[16] The numbers of year level NCEA Level 3 passes are obtained from the temporary results in the process of producing the total number of NCEA Level 3 passes.

school NCEA Level 1-3 pass rate participation [17] by decile were produced (see Appendix III), which show an overall positive relationship between the NCEA Level 1-3 pass rate and the SES decile for both the test periods[18]. The boxplots for the two test periods in Figure 3.1 also depict a positive relationship between the mean NCEA pass rates participation and decile. A further review on the standard deviation of the pass rates by decile disclosed that for 2009-2011, the standard deviation of pass rate for each of decile 1-9 varies between 0.36 and 0.40 with a higher value for decile ten, being 0.43; whereas for 2004-2006, the standard deviation is in a range between 0.31 and 0.37 for decile 1-9, with a higher value for decile ten, being 0.40. The higher standard deviation for decile ten in both periods is due to the co-existence of NCEA and Cambridge International Examination (CIE) in a few decile ten schools,[19] which focus on the CIE outcomes and receive much fewer NCEA passes compared to other decile ten schools. As a consequence, the pass rate for decile ten schools has a higher variety. Nevertheless, the stable standard deviation of NCEA pass rate for decile 1-9 schools suggests an overall positive relationship between NCEA pass rate and SES decile; even for decile ten, the small number of schools using CIE has little impact on this relationship.

To further investigate the importance of SES decile to school performance measurement, the relationship between the Government grant per student and SES decile is illustrated in Figure 3.2. It shows a negative relationship. Given that school exam pass rate directly relates to academic efficiency and indirectly relates to academic effectiveness, [20] whereas the Government grant per student is directly associated with cost efficiency and indirectly associated with overall efficiency,[21] the decile effect has to be adjusted for in each model of cost efficiency, academic efficiency, academic effectiveness and overall efficiency. This is consistent with the view of applying categorical DEA technique at each stage of decomposition in the network DEA model, that is, to measure a school's performance within its category in which all the groups have the same or less favorable socio-economic conditions, as discussed in chapter 2.

[17] The NCEA pass rate participating is calculated by total number of NCEA Level 1-3 passes divided by total number of exam participants. Numbers of participants are computed based on the data in 'Qualification-Statistics-School-yyyy'.

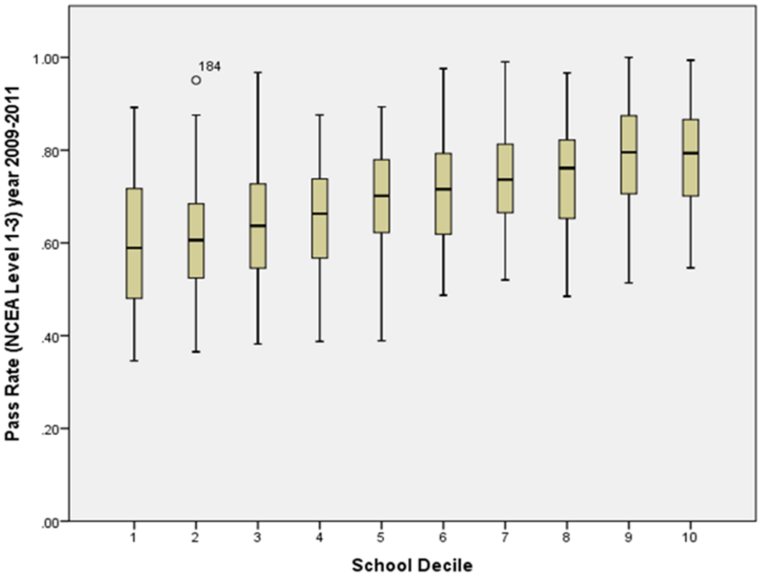
[18] In this study, two test periods refer to 2009-2011 and 2004-2006, unless otherwise stated. Test periods are discussed in section 3.3.4.

[19] For example, Macleans College and Auckland Grammar School.

[20] NZQA statistics show that the number of exam participants is significantly and positively related to school roll, therefore NCEA pass rate participation is positively related to academic efficiency. Given that school roll and NCEA qualification passes are also correlated to academic effectiveness, it is necessary to incorporate SES effect into the calculation of academic effectiveness.

[21] Since decile was designed to differentiate the allocation of Government grants, it is also related to the calculation of overall efficiency.

Figure 3.1 Decile and NCEA pass rate participation



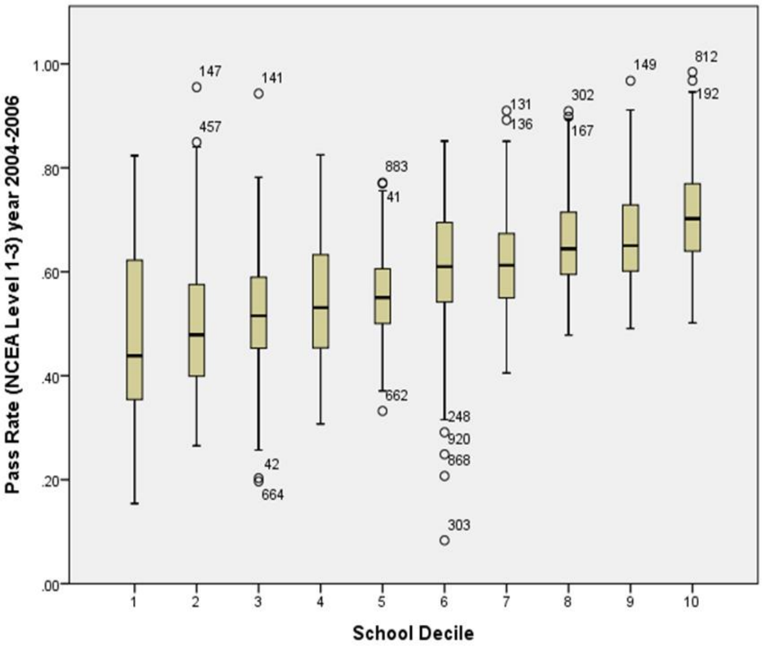
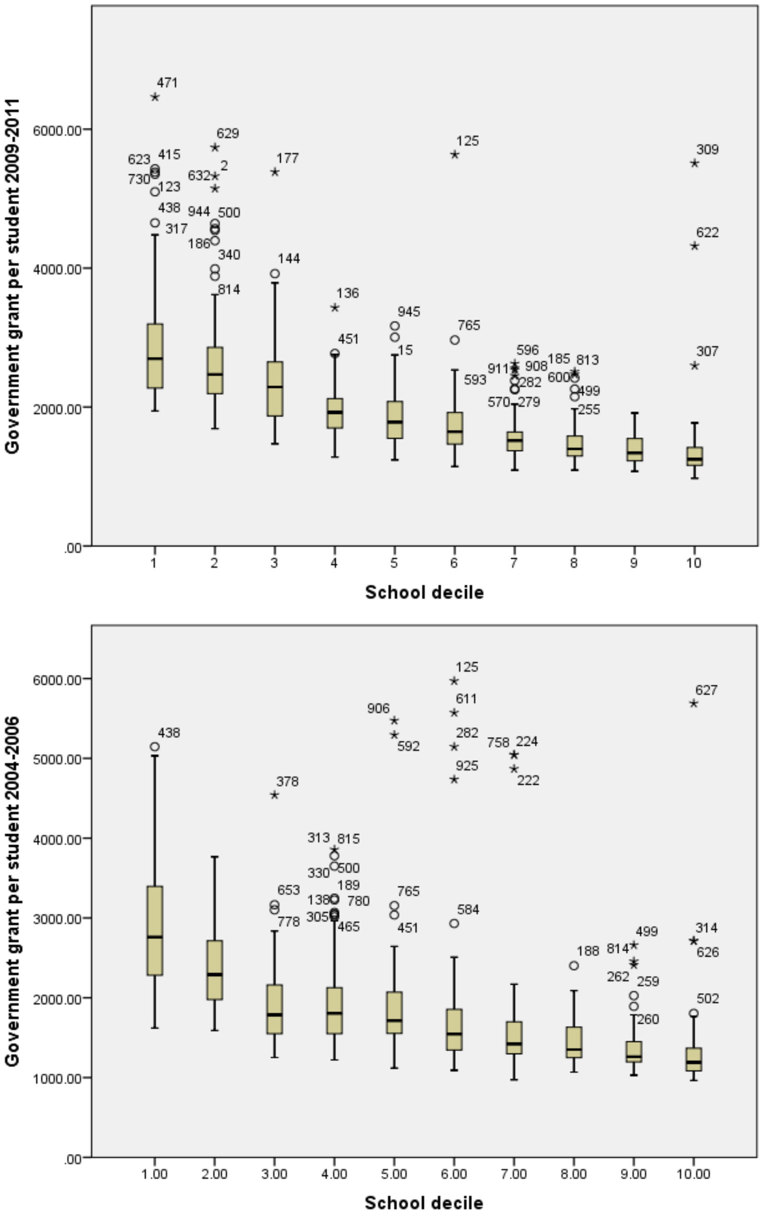


Figure 3.2 Decile and the Government grant per student



3.2.3 Other characteristics relevant to school performance

According to Alexander et. al (2010) and the NZQA statistics (New Zealand Qualifications Authority, 2013), there are a few school characteristics other than the DEA variables that need to be considered when comparing school performance:

- Ownership

  There are two types of ownership for schools in the sample: integrated and state-owned (non-integrated). Research has revealed that integrated schools are more efficient than state-owned schools in terms of academic achievements, partially due to the support from the community of interest as a continued tradition (Alexander et.al, 2010);

- Single-sex schools

  Research has also found that single-sex schools outperform co-educational schools (Harrison, 2008; Alexander, 2010). This finding is in keeping with the point of view that single-sex schools improve or retain their performance efficiency by offering specific curriculum choices to boys or girls so that they are much more efficient than co-educational schools. Research has also revealed that girl school performance is better than boy schools (Alexander et.al, 2010), therefore, the characteristic of being single boy or single girl is also recorded for further investigation;

- Mäori density and academic records

  NZQA statistics show that academic performance of Mäori students lags behind the average level (New Zealand Qualifications Authority, 2012). As such, the number of Mäori students and number of Mäori NCEA Level 1-3 qualification passes and UE passes were collected to further examine the relationship between school performance and Mäori density.

3.3 Procedures for data collection

The data for the DEA model were obtained from the MoE and the NZQA websites. The ERO reports were downloaded from the ERO website. Data were collected to cover all the state-funded schools that provide secondary education.

3.3.1 The Ministry of Education

Two datasets were collected from the MoE and its website:

- School financial data

  All financial data from 2004 to 2011 were received from the manager of the MoE Schooling Analysis Department via emails. The datasets were sorted at year level and organized in Excel format, with the reporting notes provided. These data were initially extracted from school audited annual reports.

- School rolls

  School rolls as at 1 July from 2004 to 2011 were downloaded from the MoE's official website, known as "education counts". School SES deciles are included in the roll dataset.

3.3.2 The New Zealand Qualifications Authority

All the data in respect of school NCEA qualifications, endorsements and scholarships between 2004 and 2011 were downloaded from the NZQA official website.

3.3.3 The Education Review Office

The ERO reports for state-funded secondary education providers between 2004 and 2011 were downloaded from the ERO official website.

3.3.4 Test periods

Two test periods from 2004 to 2006 and 2009 to 2011 were determined for four reasons. First, the MoE could only provide the data up to 2012 when the data were collected. However, there were 68 secondary/expanded schools (15.5% of the total schools providing secondary education) lacking 2012 financial data. Therefore, 2011 was the latest year for the complete data in hands. Second, research question two looks at the change in school performance since NCEA was fully introduced in 2004. Thus, 2004 is chosen as the earliest year for the first test period. Third, 2004-2006 and 2009-2011 were selected for the purpose to compare the impact of the NCEA system on school performance between the two periods, given that the earlier period relates to the initial impact of the new system whereas the latter period represents the system once well-established. Finally, to ensure a large sample for DEA analysis in each test period, pooled school years[22] were included in each sample and at least two years' continuity must be guaranteed to avoid any technical distortion on the DEA results. Consequently, three continuous years were included in each test period, being 2004 to 2006 and 2009 to 2011. The concept of 'school year' and data continuity will be addressed in section 3.4.

3.4 Sample selection

To ensure each sample size is large enough for DEA analysis, the concept of "school year" is introduced to represent an individual school in a specific year within each test period. For example, Avondale College in 2010 is a school year while Avondale College in 2011 is another school year. As such, each DMU in any of the efficiency models corresponds to a school year.

[22] A school year represents an individual school in a specific year.

The target population in this study is defined as New Zealand public secondary schools that do not provide primary education. Therefore, it consists of both expanded and pure secondary school years. Two samples were formed from the target population: the full sample and the pure secondary school sample. The full sample includes both expanded school years and secondary school years whilst the pure secondary school sample contains secondary school years only. The reason for forming two samples is, the expenditure splitting for the expanded schools might cause distortion on the DEA results for schools in the full sample, whilst the pure secondary school sample may not be able to represent the target population, given that the expanded schools contribute one third of the school population. Accordingly, it is necessary to calculate the efficiencies for both the samples and compare the results between them.

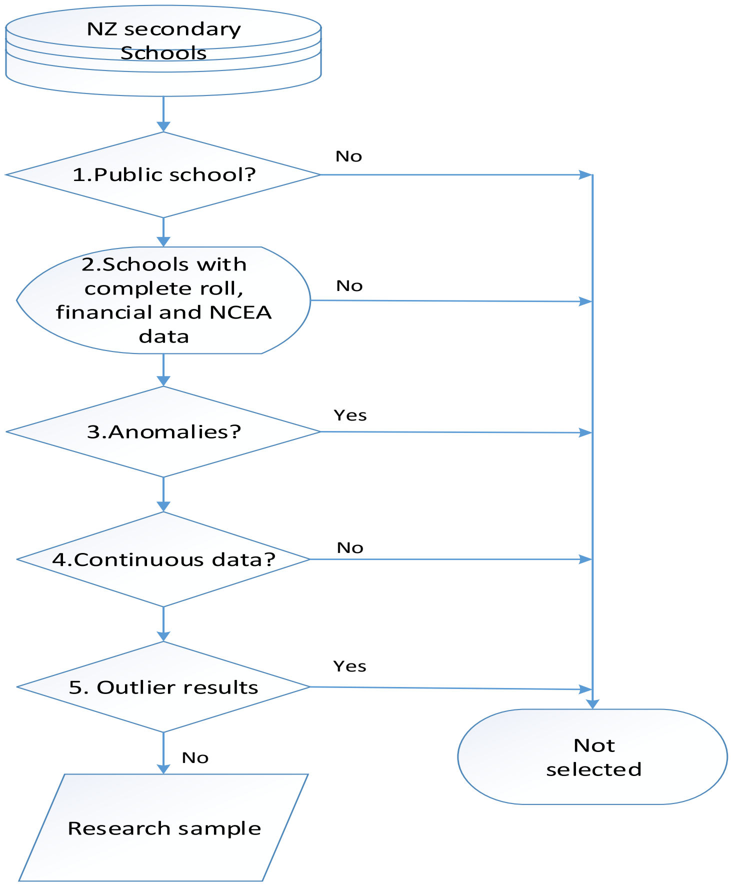
As a starting point for sampling, schools providing primary education were excluded from the target population. Therefore, it is unnecessary to set a filter to remove composite schools from the samples.

3.4.1 Filters and data validation

As shown in Figure 3.3, five filters were set to exclude certain schools[23] from each research sample: private schools, schools without complete roll, financial data and NCEA data, schools with anomalies, schools without continuous data and schools with outlier results.

[23] When a school is filtered out, its related school years are all excluded from the sample.

Figure 3.3 Procedures of sample selection



Adapted from Harrison (2008, p. 116)

The reasons for setting these filters are listed below.

First, the DEA model was developed to measure the performance of public schools which provide secondary education. Therefore, private schools were excluded from the samples.

Second, schools without roll data, teachers' salaries, operational expenditures or NCEA data were removed from the samples.

Third, given that all the data were obtained from secondary sources, a validation procedure was set to identify anomalies, being the schools with extreme cost per student. Stem-and-Leaf plotting was used to filter out these anomalies. An additional review showed that most of the anomalies were in very small schools (roll<300), and their high cost per student was mainly caused by annual fixed costs; the others were under additional construction or just opened during the test periods. These schools were removed from the samples as their extreme cost per student may distort the productivity frontier in either cost efficiency or overall efficiency model.

Fourth, schools without continuous data by years across each test period could distort the continuity or convexity of the frontier in a DEA model using pooled school data.[24] Therefore, schools without two years' continuous data in each test period were excluded from the samples.

Finally, all zero results for NCEA Level 1-3 qualification passes or number of UE passes and number of NCEA L3 M&E may cause an outlier DEA efficiency score, which is not a visible figure[25] to be used for data analysis. Therefore, schools with all zero number of NCEA passes or all zero number of UE and number of NCEA L3 M&E were excluded from the samples.

Tables 3.3 and 3.4 display the summaries of how research samples were formed for the two test periods of 2009-2011 and 2004-2006.

[24] In this study, pooled data refers to school data in different years within each test period, also called school year data.
[25] The software reports the result as 'Outbound'.

Table 3.3 Sample summary for the period 2009-2011

Data Items	Number of School Years (2009-2011)	Percentage of Each Data Item
Total school years of NZ secondary schools	339+ 343+ 345=1027	100%
Exclusions:		
1. Private secondary schools	21+22+22=65	6.33%
2. Lack of data	7+9+8=27	2.63%
2.1 No financial data	1+5+4=10	0.97%
2.2 No NCEA data	3+4+4=11	1.07%
2.3 No teachers' salaries	3	0.27%
Total excluded school years before splitting expenditures for expanded secondary schools (year 7-15)	65+27=92	8.96%
Total school years left before splitting expenditures for expanded secondary schools	1027-92=935	91.04%

	Full sample Total secondary school years=1027	Pure secondary schools (years 9-15) Total pure secondary school years =235+237+239=711
Number of secondary school years	311+312+315=938	219+219+222=660
Anomalies	18+21+14=53	14+13+14=41
Not continuous for 2 years	17	4
Outlier results	3	0
Number of school years in sample	865	615
Percentage of sample school years	**84.23%** (see note)	**86.50%** (see note)

Note: the percentage of sample school years for the full sample is the proportion of sample school years in the population; the percentage of sample school years for the pure secondary school sample is the proportion of sample school years in all the pure secondary school years.

Table 3.4 Sample summary for the period 2004-2006

Data Items	**Number of School Years (2009-2011)**	**Percentage of Each Data Item**
Total school years of NZ secondary schools	338+ 329+ 335=1002	100%
Exclusions:		
1. Private secondary schools	21+13+18=52	5.19%
2. Lack of data	14	1.40%
2.1 No financial data	0	0%
2.2 No NCEA data	5+6+3=14	1.40%
2.3 No teachers' salaries	0	0%
Total excluded school years before splitting expenditures for expanded secondary schools (years 7-15)	26+19+21=66	6.59%
Total school years left before splitting expenditures for expanded secondary schools	1002-66=936	93.41%

Full sample	**Total secondary school years=1002**	**Pure secondary schools (years 9-15) Total pure secondary school years =243+236+233=712**
Number of secondary school years	312+310+314=936	224+219+220=663
Anomalies	16+18+19=53	12+11+14=37
Not continuous for 2 years	7+1+6=14	5+1+4=10
Outlier results	4+3+6=13	2+2+1=5
Number of school years in sample	856	611
Percentage of sample school years	**85.43%** (see note)	**85.81%** (see note)

Note: the percentage of sample school years for the full sample is the proportion of sample school years in the population; the percentage of sample school years for the pure secondary school sample is the proportion of sample school years in all the pure secondary school years.

3.4.2 Comments on the datasets

The datasets are suitable for DEA analysis due to the following characteristics:

- The population is defined for this study: it contains all the public expanded (years 7-15) and pure secondary (years 9-15) schools in New Zealand. Two samples were formed: the full sample and the pure secondary school sample. For the period 2009-2011, the full sample contains 84.23% school years of the population whereas the pure secondary school sample includes 86.5% school years of the total pure secondary schools (for the period 2004-2006, full sample: 85.43%; pure secondary sample: 85.81%). Therefore, it is considered that the full sample presents an inclusive representation of the target population whereas the pure secondary school sample is also a good representation of the pure secondary schools in New Zealand.
- For each test period, the full sample contains more than 800 school years (DMUs) while the pure secondary school sample includes more than 600 school years. As such, each sample size is technically sufficient for DEA analysis.
- The homogeneity of sample schools is high in terms of the educational process at each school.
- There is a high degree of data validity within each sample because all the data were obtained from the MoE and the NZQA, who administered and collected these data through rigorous and consistent procedures. For example, school financial data were extracted from school annual financial reports which are prepared as per the *Generally Accepted Accounting Practice* (GAAP) as well as the MoE's accounting guidelines, and have been audited by independent auditors; the academic outcome data are organized, monitored and modified by the NZQA. All these data are subject to scrutiny from schools and the public, given that NCEA data are promptly published on the NZQA website and can be accessed for a long period.

However, schools that opened or closed during each test period were not included in the samples due to the restriction of data continuity. This might cause a 'self-selection' issue as schools with worst performance could be excluded from the samples, because schools with deteriorating performance might be required to close whilst just opened schools may have poor performance at the beginning due to their large expenditures and small sizes.

3.5 Summary

This chapter defines all the variables for the network DEA model: teachers' salaries and other operational expenditures, the numbers of NCEA Level 1-3 qualification passes, roll, the number of UE and number of NCEA L3 M&E. As a non-discretionary variable, school SES decile is incorporated in each categorical DEA model to calculate the cost efficiency, academic efficiency, academic effectiveness and overall efficiency.

In order to better represent New Zealand secondary schools and measure school performance more accurately, two samples were formed for each test period, being the full sample (combination of the pure secondary schools and the expanded secondary schools), and the pure secondary school sample. The expenditures for intermediate education in the expanded secondary schools were removed to ensure that all the DEA variables are related to secondary education only. Moreover, to minimize any possible distortion caused by inflation effect, expenditures were all indexed using the GNE deflators.

Data were collected from the MoE, the NZQA and the ERO. Two test periods of 2004-2006 and 2009-2011 were chosen to compare the different impact of NCEA on school performance between the two periods: just after NCEA's introduction and more recently once NCEA had become well-established.

The major advantage of the samples is that they are technically suitable for a DEA analysis in terms of sample size, homogeneity and data validity. The limitation of the datasets is that some schools with the worst performance might be excluded.

Chapter 4 presents and discusses the DEA results for New Zealand secondary school performance.

Chapter 4 Data Analysis and Discussion

4.1 Introduction

In this chapter, New Zealand school performance under the NCEA system is analyzed using the efficiencies produced by the network DEA model. Research question two and three are answered based on the DEA results, Malmquist indices, regression results and ERO reports for the top ten and bottom ten schools identified by the performance measurement system.

This chapter is structured as follows: after the introduction in section 4.1, sections 4.2 and 4.3 describe and compare the major characteristics of the two samples. Section 4.4 provides descriptive statistics of major school characteristics for the two samples. Section 4.5 discusses the DEA efficiencies and examines changes in school performance since the introduction of NCEA. Section 4.6 examines the relationships between school performance and school characteristics that are not included in the network DEA model. Section 4.7 analyses the possible reasons for good and poor school performance based on the ERO reports for the top ten and bottom ten schools identified by the performance measurement system. Section 4.8 is the section summary. Of note, section 4.5.4 and section 4.6.2 answer research question two while section 4.7.2 addresses research question three.

4.2 Description of samples

As discussed in chapter 3, two samples were selected in each test period and DEA was used to measure performance in terms of cost efficiency, academic efficiency, academic effectiveness and overall efficiency. Each sample contains pooled data using the ‘school year’ as the unit of analysis, which represents any information of interest for a specific school in a specific year.

The full sample contains 291 schools and 865 school years during 2009-2011, and contains 288 schools and 856 school years during 2004-2006. It therefore has 2.97 observations per school during each test period. The pure secondary school sample contains 206 schools and 615 school years during 2009-2011, and contains 207 schools and 611 school years during 2004-2006. Thus it has 2.98 observations per school during 2009-2011 and 2.95 observations per school during 2004-2006.

4.3 Comparison of key characteristics between samples

The key characteristics of the full sample and the pure secondary school sample during 2009-2011 and 2004-2006 are displayed in Tables 4.1 and 4.2. Schools that closed or opened during each test period were removed due to a lack of data continuity.

The discussion below for the comparison between the full sample and the pure secondary school sample is based on the data for the period 2009-2011. The corresponding statistics of the period 2004-2006 are included in brackets.

It appears that the pure secondary school sample for the two test periods is more homogeneous than the full sample which includes 27.4% schools providing intermediate education (2004-2006: 26.75%). The differences in ownership variable are significant between the two samples: with the pure secondary school sample including 9.76% (2004-2006: 10.48%) of state-integrated schools whereas 20.12% (2004-2006: 21.26%) of the full sample are state-integrated schools. There is a slight decrease in the percentage of state-integrated schools in each sample from the earlier to later test period.

The pure secondary school sample consists of significantly larger schools ($p<0.001$), with a mean roll of 990 (2004-2006: 976) compared to 887 (2004-2006: 864) for the full sample. As a consequence, significant differences are revealed for both total operating expenditures and cost per student, indicating that secondary education has larger rolls and higher total costs compared to intermediate education. However, the operating cost per student is significantly lower for the pure secondary school sample with $8,057 (2004-2006: $6,174) compared to $8,356 (2004-2006: $6,334) for the full sample. This feature is evidenced by both the total operating expenditure variables and the adjusted expenditure variables that exclude the spending for intermediate education as detailed in chapter 3, and reflects the larger roll size of schools in the pure secondary school sample.

There are no significant differences in the density of single sex/ co-educational schools within either the pure secondary school sample or the full sample, with 29.6% (2004-2006: 31.11%) of single sex schools in the pure secondary school sample compared to 29.71% (2004-2006: 31.20%) for the full sample.

Given there are some differences in characteristics between the pure secondary school sample and the full sample, it is expected that the pure secondary schools will have higher mean cost efficiency score than the full school sample due to the lower operating cost per student. Further, it is also expected that the variation in any type of the efficiency scores for the pure secondary school sample would be lower due to their more homogeneous nature.

Table 4.1 Characteristics of sample for the period 2009-2011

Characteristics	Full sample	Sample	Pure secondary school sample	Sample	Independent T-test (difference of means)[26]
Sample size					
Number of schools	291		206		
Number of school years	865		615		
Mean school years per school	2.97		2.98		
School ownership					
State-owned	691	79.88%	555	90.24%	
State-integrated	174	20.12%	60	9.76%	
School type					
Expended schools yr7-15	237	27.4%	0	0	
Secondary schools yr9-15	628	82.6%	615	100%	
Gender					
Co-educated	608	70.29%	433	70.4%	
Boys	117	13.53%	80	13%	
Girls	140	16.18%	102	16.6%	
School expenditure					
Mean indexed total expenditure	6,745,718	NA	7,389,141	NA	-3.779**a
Mean indexed cost per student	8,356	NA	8,057	NA	2.826**a
Mean indexed adjusted operating expenditure (secondary cost only)	6,406,357	NA	7,389,141	NA	-5.65***b
Mean indexed adjusted operating cost per student (secondary cost only)	8,417	NA	8,057	NA	4.37***b
School size					
Mean roll	887	NA	990	NA	-3.65***a
Mean secondary level roll	841	NA	990	NA	-5.21***b
SES Decile					
Mean decile	5.69	NA	5.52	NA	1.25a

Significance: ***p <0.001, **p<0.01 (2- tailed); effective size[27] (r), a=0.01<r<0.1, b=0.1<r<0.2.

[26] t values for each mean compare are listed in this column, followed by the significance.

Table 4.2 Characteristics of sample for the period 2004-2006

Characteristics	Full sample	Sample	Pure secondary school sample	Sample	Independent T-test (difference of means) [28]
Sample size					
Number of schools	288		207		
Number of school years	856		611		
Mean school years per school	2.97		2.95		
School ownership					
State-owned	674	78.74%	547	89.52%	
State-integrated	182	21.26%	64	10.48%	
School type					
Expended schools yr7-15	229	26.75%	0	0	
Secondary schools yr9-15	627	73.25%	611	100%	
Gender					
Co-educated	589	68.8%	427	69.89%	
Boys	122	14.25%	84	13.75%	
Girls	145	16.95%	100	16.36%	
School expenditure					
Mean indexed total expenditure	5,095,338	NA	5,673,407	NA	-4.31***b
Mean indexed cost per student	6,334	NA	6,174	NA	2.97**a
Mean indexed adjusted operating expenditure (secondary cost only)	4,872,806	NA	5,673,407	NA	-5.84***b
Mean indexed adjusted operating cost per student (secondary cost only)	6,478	NA	6,174	NA	5.28***b
School size					
Mean roll	864	NA	976	NA	-4.05***b
Mean secondary level roll	820	NA	976	NA	-5.54***b
SES Decile					
Mean decile	5.77	NA	5.64	NA	0.93a

Significance: ***p <0.001, **p<0.01 (2- tailed); effect size[29] (r), a=0.01<r<0.1, b=0.1<r<0.2.

[27] Effective size for T-test, $r=\sqrt{\frac{t^2}{t^2+df}}$

[28] Ibid.

[29] Effect size is a measure of how important that the significant differences between two practical terms are (Field, 2013, p. 376). Effect size for T-test, $r=\sqrt{\frac{t^2}{t^2+df}}$

4.4 Descriptive statistics

4.4.1 Sample statistics

The descriptive statistics for the model variables are provided in Tables 4.3 and 4.4, along with the results of *t*-tests for comparisons of the means between the full sample and the pure secondary school sample. Given that all the variables do not conform to a normal distribution, a non-parametric Mann-Whitney U-Test is also conducted.

As discussed earlier, the distribution of teachers' salaries and other resources are closely associated with school roll, this relationship is reflected by the wide range of each, along with the considerable variation of school roll. The mean differences of school roll between the full sample and the pure secondary school sample are significant during both test periods.

Significant differences were identified for all the expenditure variables, the numbers of NCEA Level 1-3 passes and the number of UE qualifications between the full sample and the pure secondary school sample in both test periods. The difference in the number of NCEA L3 M&E is less significant (2009-2011: not significant; 2004-2006, significant at 0.05 level). This is because NCEA Level 3 high level endorsements are mainly achieved by pure secondary school students, while the pure secondary schools form a subset of the full sample.

Based on the significant negative *t*-values of the expenditures, NCEA passes and UE, the pure secondary school sample has overall higher level expenditures, more NCEA passes, and more UE passes than the full sample. This feature is also related to the fact that the pure secondary school sample contains larger schools which are operated using higher level expenditures and have more students passing NCEA due to their bigger sizes.

The Mann-Whitney U-test provided similar results to the T-test in terms of significance levels of the difference in variable means between the full sample and the pure secondary school sample. Z values of the U-test for the model variables are reported together with the effect size[30] levels indicated in Tables 4.3 and 4.4.

[30] Effect size for U-test is denoted by r, r=$\frac{z}{\sqrt{N}}$

Table 4.3 Descriptive statistics- variables in network DEA model for 2009-2011

	Full sample (n=865)				Pure secondary school sample (n=615)				Independent T-test[31]	Mann-Whitney U test[32]
	Min	Max	Mean	Std. Dev	Min	Max	Mean	Std. Dev		
SES decile	1	10	5.69	2.55	1	10	5.52	2.53	1.25a	-1.219a
Indexed teachers' salaries	89,376	13,086,800	4,054,814	2,099,730	89,376	13,086,800	4,685,385	2,129,390	-5.66***b	6.644***b
Indexed other expenditures	472,057	8,734,716	2,351,543	1,259,029	588,073	8,734,716	2,703,756	1,268,228	-5.29***b	6.449***b
Roll (secondary level 9-15)	120	3,046	841	540	184	3,046	990	554	-5.2***b	6.040***b
Number of NCEA Level 1 passes	16	640	151	100	26	640	177	104	-4.76***b	5.423***b
Number of NCEA Level 2 passes	11	613	130	91	26	613	152	96	-4.52***b	5.146***b
Number of NCEA Level 3 passes	1	453	73	61	4	453	84	66	-3.32***a	3.666***a
Number of university entrance	1	431	65	58	3	431	75	64	-3.12**a	3.387**a
Number of NCEA Level 3 Merit and Excellence endorsements	0	220	20	26	0	220	23	29	-1.89a	2.009*a

Significance: ****p* <0.001, ***p*<0.01, * *p*<0.05 (2- tailed); effect size: a=r<0.1, b=0.1<*r*<0.2

[31] This column records *t*-value.
[32] This column records *z* values for the U-test.

Table 4.4 Descriptive statistics- variables in network DEA model for 2004-2006

	Full sample (n=856)				Pure secondary school sample (n=611)				Independent T-test[33]	Mann-Whitney U test[34]
	Min	**Max**	**Mean**	**Std. Dev**	**Min**	**Max**	**Mean**	**Std. Dev**		
SES decile	1	10	5.77	2.60	1	10	5.64	2.53	0.93a	-0.939a
Indexed teachers' salaries	185,508	13,248,849	3,652,524	1,966,280	185,508	13,248,849	4,255,695	1,973,906	-5.78***b	6.857***b
Indexed other expenditures	259,526	4,147,795	1,220,281	695,792	348,537	4,147,795	1,417,712	702,143	-5.34***b	6.482***b
Roll (secondary level 9-15)	129	3,097	820	527	180	3,097	976	533	-5.54***b	6.557***b
Number of NCEA Level 1 passes	10	695	138	97	21	695	162	101	-4.63***b	5.574***b
Number of NCEA Level 2 passes	9	596	110	87	17	596	130	94	-4.04***b	4.908***b
Number of NCEA Level 3 passes	1	398	55	54	2	398	65	59	-3.08**a	3.678***a
Number of university entrance	1	392	52	51	2	392	60	56	-2.97**a	3.548***a
Number of NCEA Level 3 Merit and Excellence endorsements	0	153	14	20	0	153	16	22	-2.13*a	2.627**a

Significance: ****p* <0.001, ***p*<0.01, * *p*<0.05 (2- tailed); effect size: a=$r<0.1$, b=$0.1<r<0.2$

[33] This column records t-value.
[34] This column records *z* values for the U-test.

4.4.2 Normality test for samples

The distribution details of the samples by school size/roll, in conjunction with the mean SES decile in the two test periods, are displayed in Tables 4.5 and 4.6.

Both tables show a large range in school size and a large percentage of schools in each sample with rolls smaller than the mean value (2009-2011: FS[35]=887, PS[36]=990; 2004-2006: FS=864, PS=976), whereas a much smaller proportion of large schools. Moreover, the two tables also reveal that school size is positively associated with the mean SES decile, i.e., large schools generally have higher decile than small schools.

It appears that the distribution of the school rolls is not symmetrically centred on the mean roll therefore the variables of school rolls in the two samples are not normally distributed (positive skew and long right tail). To determine whether non-parametric techniques are needed for data analysis, statistical tests were conducted on all the model variables. Appendix IV demonstrates the skewness and kurtosis statistics for all the DEA variables in the two test periods.

According to Appendix IV, all variables, apart from the SES decile, have skewness to some extent with most greater than one. This indicates that all of them do not conform to a normal distribution. Further, most variables have positive skewness because they are positively related to school size which is also positively skewed. Similarly, all variables except for the SES decile have positive kurtosis showing that they are not normally distributed. Most variables have kurtosis in excess of one with school academic variables having particularly high kurtosis values. This is also consistent with the view that school academic outputs and outcomes are positively related to school roll through the SES effect, given that the majority of large schools have high SES decile.

It is evident that the pure secondary school sample in both test periods has lower skewness and higher kurtosis. This indicates that there is a greater proportion of schools in the pure secondary school sample distributed in the right area (greater than mean) and the distribution is less flat than the full sample.

It is sensible that SES decile has a low skewness and a negative kurtosis around -1 due to its construction as a decile. It has a much flatter distribution in both samples compared to other variables.

---

[35] FS refers to the full sample.
[36] PS denotes the pure secondary school sample.

In addition to the skewness and kurtosis tests, normal Q-Q plots, histograms and stem-and-leaf plots on each model variable were examined (not reported here) and all these tests showed that the variables are not normally distributed.

Accordingly, a non-parametric statistical approach is required to conduct data analysis, together with parametric tests.

Table 4.5 School distribution by size for the period 2009-2011

School roll	Full sample (mean roll=887[37])			Pure secondary school sample (mean roll=990)		
	Number of school years	Percentage of sample	SES mean	Number of school years	Percentage of sample	SES mean
<=300	58	6.7	4.29	22	3.58	3.68
301-600	249	28.79	4.84	139	22.6	4.2
601-900	221	25.55	5.45	168	27.32	4.98
901-1,200	133	15.38	6.38	97	15.77	6.24
1,201-1,500	90	10.40	6.38	81	13.17	6.17
1,501-1,800	52	6.01	7.4	50	9.76	7.34
1,800-2,100	30	3.47	7.17	26	4.23	7
2,101-2,400	16	1.85	6.88	16	2.6	6.88
2,401+	16	1.85	8.13	16	2.6	8.13
Total	865	100	5.69	615	100	5.52

[37] Roll number here is the actual roll of the whole school rather than the roll for secondary students (years 9-15).

Table 4.6 School distribution by size for the period 2004-2006

School roll	Full sample (mean roll=864[38])			Pure secondary school sample (mean roll=976)		
	Number of school years	Percentage of sample	SES mean	Number of school years	Percentage of sample	SES mean
<=300	58	6.78	4.90	18	2.95	4.11
301-600	251	29.32	5.02	144	23.57	4.34
601-900	208	24.30	5.49	150	24.55	5.31
901-1,200	159	18.57	6.28	126	20.62	5.99
1,201-1,500	93	10.86	6.35	90	14.73	6.3
1,501-1,800	35	4.09	7.46	32	5.24	7.38
1,800-2,100	20	2.34	7.5	19	3.11	7.47
2,101-2,400	18	2.1	7.5	18	2.94	7.5
2,401+	14	1.64	8.43	14	2.29	8.43
Total	856	100	5.77	611	100	5.64

4.4.3 Correlation

Non-parametric correlations (i.e. Spearman's Rho) were computed for model variables in both the full sample and the pure secondary school sample in both test periods (see Tables 4.7 and 4.8), the expenditure variables for both test periods are GNE indexed.

The correlation results show that the SES decile is positively related to expenditure variables, roll and academic measures with weaker correlation to expenditures (FS: $r<0.3$, $p<0.001$; PS: $r<0.44$, $p<0.001$) and stronger to academic indicators (FS: $r\geq3.98$, $p<0.001$; PS: $r\geq0.5$). The SES is correlated to the roll at a moderate level (FS: $0.25< r<0.32$, $p<0.001$; PS: $0.4< r<0.45$, $p<0.001$).

Other model variables are strongly and positively correlated to each other, and all significant at 0.001 level. It might be argued that in a DEA model, a strong correlation between two input/output variables affects the discrimination of DEA analysis because a strong relationship between two measures may indicate that these measures reflect the same thing (Banxia Software Limited, 2010). However, in this study, each measure in any of the sub DEA models represents different aspects of the underlying phenomena, the significant correlations between some inputs or outputs exist because they are all positively related to school roll, or they have inherent relationships. For example,

[38] ibid

teachers' salaries and other resources have a strong correlation ($r=0.911$, $p<0.001$) as they are all positively related to roll, but they measure different school resources therefore they are included as the inputs in cost efficiency model; the number of NCEA Level 1 passes is significantly correlated to the number of NCEA Level 2 passes ($r=0.967$; $p<0.001$) due to their inherent relation via the credit requirement (see section 2.2.2.2). However, they measure different levels of NCEA qualifications thus both of them are selected to measure the academic efficiency and academic effectiveness.

On the other hand, it is expected that inputs and outputs in a DEA model should have strong correlations to ensure the validity of measurement. In this regard, the overall strong correlations (FS: $r>0.59$, $p<0.001$; PS: $r>0.61$, $p<0.001$) between inputs and outputs in all the sub-models (i.e., cost efficiency, academic efficiency and academic effectiveness models), satisfy this rule.

It is apparent that the correlations between variables in the pure secondary school sample are overall stronger than the full sample. This is consistent with the more homogenous nature of the pure secondary school sample and schools are more symmetrically distributed towards the mean point with overall higher possibilities (lower skewness and higher kurtosis shown in Appendix IV). As a result, stronger positive relationships between model variables exist in the pure secondary school sample, compared with the full sample.

Table 4.7 Spearman's correlations for DEA model variables for 2009-2011[39]

	SES	InxTeacher$	InxOther$	Roll	#L1	#L2	#L3	#UE	#L3M&E
Full sample n=865									
SES	1								
InxTeacher$	.265	1							
InxOther$	.263	.911	1						
Roll	.316	.977	.916	1					
#L1	.398	.933	.868	.963	1				
#L2	.445	.925	.863	.955	.967	1			
#L3	.606	.812	.754	.849	.888	.920	1		
#UE	.651	.793	.744	.833	.870	.901	.984	1	
#L3M&E	.707	.637	.598	.678	.746	.786	.899	.913	1
Pure secondary school sample n=615									
SES	1								
InxTeacher$	.396	1							
InxOther$	.412	.899	1						
Roll	.443	.980	.918	1					
#L1	.518	.923	.861	.951	1				
#L2	.564	.922	.867	.948	.969	1			
#L3	.680	.828	.786	.858	.898	.931	1		
#UE	.710	.815	.780	.847	.881	.913	.984	1	
#L3M&E	.749	.690	.671	.718	.784	.821	.904	.917	1

Note: all correlations are significant at 0.001 level (2-tailed).

Table 4.8 Spearman's correlations for DEA model variables for 2004-2006

	SES	InxTeacher$	InxOther$	Roll	#L1	#L2	#L3	#UE	#L3M&E
Full sample n=856									
SES	1								
InxTeacher$	.204	1							
InxOther$	.160	.887	1						
Roll	.255	.976	.901	1					
#L1	.411	.902	.827	.941	1				
#L2	.489	.862	.798	.905	.949	1			
#L3	.619	.747	.693	.794	.875	.920	1		
#UE	.639	.731	.675	.781	.870	.913	.989	1	
# L3M&E	.654	.645	.594	.693	.791	.836	.915	.925	1
Pure secondary school sample n=611									
SES	1								
InxTeacher$	.348	1							
InxOther$	.273	.863	1						
Roll	.404	.976	.877	1					
#L1	.552	.889	.792	.931	1				
#L2	.605	.862	.776	.904	.949	1			
#L3	.694	.775	.692	.818	.892	.934	1		
#UE	.709	.755	.669	.800	.888	.927	.989	1	
# L3M&E	.703	.689	.617	.733	.828	.861	.924	.935	1

Note: all correlations are significant at 0.001 level (2-tailed).

[39] In this thesis, "#" denotes 'the number of' and the prefix 'Inx' means indexed using GNE deflators in all the tables unless otherwise stated.

4.4.4 Comparison of model variable means across school years

Since each sample contains pooled data for the years during 2004-2006 or 2009-2011, and model variables do not conform to a normal distribution, both parametric (one-way ANOVA) and non-parametric approaches (Kruskal-Wallis test) were applied to compare the means of all these variables between years in each test period.

Both one-way ANOVA and Kruskal-Wallis tests provide similar results in terms of the distribution of each model variable across years. However, results generated from the Kruskal-Wallis test are considered more robust due to the non-normality of the research samples. Apart from the SES, roll and the number of NCEA Level 1 passes, all variables (i.e. expenditures and academic measures) vary significantly across years. In particular, they have significant differences ($p<0.001$) between 2004 and 2011. However, each of them has no significant variation within each test period.

It is noteworthy that the variable of other operating expenditures varies significantly between the earlier test period (2004-2006) and later test period (2009-2011). As such a further investigation was conducted (Appendix V Table A5.1). Results show that school total property expenditure has a significant increase from the earlier period to the later period, and this term has the highest effect size compared to the total administration expenditure and total learning expenditure. Given the dollar variables in Appendix V are indexed thus significant difference represents an absolute increase in spending. The primary reason for this variation is the change of accounting policy for school annual reporting: since 2007, all schools are required to record the expenditure for the 'use of land and buildings' (UoLB) under the account of 'property expenses' (Ministry of Education, 2016). Furthermore, the significant increase in the 'costs of heat, light and water' under 'property expense' account is the secondary reason for the difference as energy prices grew considerably from 2004 to 2011 (see Table A5.2 in Appendix V).

Nevertheless, given the change in accounting policy since 2007, an issue is whether the amount of other expenditures needs to be adjusted as an input for both the cost efficiency and overall efficiency models. Because school years are classified into 2009-2011 and 2004-2006 groups for independent DEA analysis and the change in accounting policy only applies to 2009-2011 group, no adjustment could be made for 2004-2006 group due to the absence of UoLB data for this period. As such, the decisive factor in this issue is whether the amount of UoLB should be removed from other operational expenditures in the DEA models for 2009-2011 group. As discussed in section 3.2.1, inputs for cost efficiency or overall efficiency models should include all the school resources related to the education process. In this regard, the amount of UoLB should not be taken out for the calculation of efficiency scores. To examine whether there are any differences in efficiency scores under the model with UoLB included and the model without UoLB included, two *t*-tests were undertaken and the results show that there are no significant differences in both mean cost efficiency and mean overall efficiency (not report here). Accordingly, with regard to the calculation of efficiency scores, no adjustments against UoLB effect were necessary for the DEA analysis in 2009-2011 group. However, for the purpose of comparing the efficiency before and after the change in accounting policy, UoLB should be removed for the calculation of efficiency relating to the period after 2007, so that the efficiencies for comparison are produced on the same basis. This issue will be addressed again in section 4.5.4.

4.4.5 Summary and discussion on descriptive statistics

Statistical tests on model variables show that all of them do not conform to a normal distribution. This indicates that non-parametric approaches are needed for data analysis. Moreover, the strong correlations between inputs and outputs evidence the reasonableness for the selection of DEA variables.

Differences in the means of model variable exist between the full sample and the pure secondary school sample, but the impacts of these differences are small (effect size<0.2).

The comparison of variables between years shows that school expenditures and academic outputs and outcomes are significantly larger in 2011 compared to 2004, in particular, the total amount of other operational expenditures varies significantly between earlier and later test periods. A further investigation revealed that this variation is primarily due to the change in accounting policy for school financial reporting in 2007, that is, schools are required to record UoLB under the account of ‘Property Expenses’ since then. As such, for the purpose of comparing the efficiency before and after the change, UoLB should be removed from the school expenditures for the calculation of cost efficiency or overall efficiency in any years when the new accounting policy applies.

4.5 DEA results

4.5.1 Efficiency scores and school characteristics

4.5.1.1 Descriptive statistics of efficiency scores

School efficiency scores for both samples in the two test periods were computed[40] in terms of cost efficiency, academic efficiency, academic effectiveness and overall efficiency. The major difference between the two samples is that the full sample includes not only pure secondary schools (years 9-15) but also expanded secondary schools (years 7-15), for which the school expenditures and rolls have been adjusted to secondary education level only. Of note, the pure secondary school sample is not the exact subset of the full sample, given that 2.14% schools in the pure secondary school sample are not included in the full sample for the period 2009-2011 and 3% for the period 2004-2006. The differences are caused by the consequence of the exclusion rule for the sampling, as the anomalies or schools without continuous data in the pure secondary school sample are different from that in the full sample. However, the pure secondary school sample can be considered as an approximate subset of the full sample.

Results in each test period under both the categorical and standard models are listed in Tables 4.9 and 4.10. The categorical model adjusts for the SES effect whereas the standard model does not consider this effect in order to provide a baseline to compare how much adjustment has been made to efficiency scores by the categorical model. To examine scale effects, both constant returns to scale (CRS) and variable returns to scale (VRS) were applied in each model.

Given that there are two test periods for this research, in this section, results for the later period (i.e., 2009-2011) will be reported then compared with those for the earlier period (i.e., 2004-2006). Differences in findings between the two periods will be identified and discussed.

As shown in Table 4.9 concerning 2009-2011, for each efficiency (i.e., cost efficiency, academic efficiency, academic effectiveness and overall efficiency), the ranges of efficiency scores are wide (between 0.46 and 0.98[41]) for both categorical and standard DEA models in both samples. The major features relating to the ranges are:

---

[40] Efficiency scores were computed using pyDEA software which was developed by the Department of Engineering Science in the University of Auckland.

[41] Here, the range of efficiency scores does not include scale efficiency scores.

- For both samples, ranges of cost efficiency are between 0.46 and 0.59 for the different models, which are lower than that of academic efficiency and academic effectiveness sitting between 0.68 and 0.85. The ranges of overall efficiency are the highest compared to all the sub-efficiencies (i.e., cost efficiency, academic efficiency and academic effectiveness), being over 0.91; these are consistent with the fact that cost management in the New Zealand educational system is highly standardized mainly driven by school roll, whereas school academic performance varies significantly along with SES, school management and funding ability.
- For all the efficiency models, the pure secondary school sample has smaller ranges than the full sample. This is consistent with the fact that the pure secondary school sample contains schools which are more homogeneous in terms of curriculum settings and school management system;
- For all the efficiency models, ranges of scores under the categorical models are smaller than under the standard models, as a consequence that the categorical models adjust upward the scores for schools in the lowest deciles and retain the scores for the highest decile schools. However, the ranges of academic efficiency, academic effectiveness and overall efficiency under the categorical models are still unexpectedly high. A review of these results showed that some decile ten schools (e.g., Auckland Grammar School and Macleans College) received low academic efficiency because they focused on the CIE system. This resulted in the wide ranges of academic efficiency, given that the categorical models do not adjust SES effect for decile ten schools. The wide ranges of academic effectiveness and overall efficiency are caused by the great variety in the number of UE passes and number of NCEA L3 M&E within each sample. Some schools achieved zero UE or NCEA L3 M&E, thus the ranges of academic effectiveness or overall efficiency scores are wide even though they are produced from the categorical models.
- Ranges of scores under the VRS mode are smaller than CRS, as VRS adjusts for economies of scale.

It is noteworthy that the pure secondary school sample has higher efficiency scores than the full sample irrespective of the type of efficiency and returns to scale. This is because more than 97% of schools in the pure secondary school sample are included in the full sample as such, the pure secondary school sample can be considered as a subset of the full sample. Technically, a DMU in the subset could get higher DEA efficiency scores compared to the original set due to the smaller sample size.

With regard to all the efficiency models, scale efficiencies for both samples are high (minimum 0.88) and this indicates that, most secondary schools in New Zealand are operating with school rolls that approximate the most productive scale. However, the wide ranges of scale efficiencies for different models, especially, for the academic effectiveness and overall efficiency models (range>0.73), suggest the existence of scale effect. Therefore, most discussions in this study are based on the results of VRS models as VRS approach adjusts the scale effect, given that the wide range of school roll in each sample also indicates the existence of scale effect.

The major features of the results for the period 2004-2006 are very similar to those for 2009-2011 in terms of score ranges for different types of efficiency and returns to scale, high scale efficiency and higher scores for the pure secondary school sample (see Table 4.10). However, it appears that there are more fully academic efficient and effective but less cost efficient school years during 2009-2011 compared to 2004-2006. This finding is in keeping with the fact that the Government has continuously increased the grants for schools to improve their academic achievements since the introduction of NCEA. As a result, cost per student increased year by year thus school cost efficiency decreased; in contrast, school academic efficiency and academic effectiveness were improved due to increases in NCEA outputs and outcomes. Changes in school performance will be further examined in section 4.5.4.

To minimize the distortion from the size effect on cost and other variables, scores under the VRS[42] model are used as the source for the major part of data analysis.

[42] VRS model adjusts size effect on efficiency scores.

Table 4.9 Descriptive statistics for network DEA model 2009-2011

	Full sample (n=865)			Pure secondary school sample (n=615)		
	# Efficient school years	Mean (St. Dev)	Range	# Efficient school years	Mean (St. Dev)	Range
1. Cost efficiency						
Categorical Model CRS	14	0.74 (0.13)	0.59	15	0.78 (0.12)	0.49
Categorical Model VRS	40	0.78 (0.12)	0.56	31	0.85 (0.09)	0.46
Categorical Model scale efficiency	NA	0.94 (0.09)	0.57	NA	0.91 (0.09)	0.45
Standard-CRS	5	0.71 (0.13)	0.59	4	0.73 (0.12)	0.53
Standard-VRS	14	0.74 (0.12)	0.57	15	0.81 (0.09)	0.46
Standard-scale efficiency	NA	0.96 (0.07)	0.57	NA	0.90 (0.09)	0.51
2. Academic efficiency						
Categorical Model CRS	19	0.79 (0.11)	0.73	22	0.81 (0.12)	0.72
Categorical Model VRS	83	0.84 (0.12)	0.70	68	0.85 (0.12)	0.70
Categorical Model scale efficiency	NA	0.97 (0.05)	0.38	NA	0.95 (0.05)	0.33
Standard-CRS	7	0.75 (0.12)	0.75	5	0.75 (0.12)	0.74
Standard-VRS	23	0.77 (0.13)	0.73	14	0.77 (0.13)	0.73
Standard-scale efficiency	NA	0.93 (0.03)	0.24	NA	0.97 (0.03)	0.15
3. Academic effectiveness						
Categorical Model CRS	49	0.72 (0.15)	0.73	39	0.81 (0.13)	0.69
Categorical Model VRS	94	0.75 (0.16)	0.72	78	0.85 (0.13)	0.68
Categorical Model scale efficiency	NA	0.96 (0.06)	0.73	NA	0.96 (0.05)	0.39
Standard-CRS	13	0.61 (0.16)	0.85	6	0.73 (0.15)	0.81
Standard-VRS	29	0.64 (0.17)	0.83	22	0.77 (0.15)	0.78
Standard-scale efficiency	NA	0.96 (0.06)	0.85	NA	0.96 (0.06)	0.69
4. Overall efficiency						
Categorical Model CRS	25	0.52 (0.21)	0.96	28	0.56 (0.22)	0.92
Categorical Model VRS	70	0.59 (0.22)	0.94	49	0.61 (0.22)	0.91
Categorical Model scale efficiency	NA	0.88 (0.14)	0.92	NA	0.92 (0.11)	0.82
Standard-CRS	5	0.39 (0.21)	0.98	7	0.39 (0.21)	0.97
Standard-VRS	19	0.42 (0.22)	0.97	14	0.43 (0.22)	0.97
Standard-scale efficiency	NA	0.93 (0.11)	0.90	NA	0.91 (0.11)	0.74

Table 4.10 Descriptive statistics for network DEA model 2004-2006

	Full sample (n=856)			**Pure secondary school sample (n=611)**		
	# Efficient school years	**Mean (St. Dev)**	**Range**	**# Efficient school years**	**Mean (St. Dev)**	**Range**
1.Cost efficiency						
Categorical Model CRS	19	0.78 (0.13)	0.57	17	0.81 (0.11)	0.48
Categorical Model VRS	58	0.82 (0.11)	0.50	42	0.85 (0.09)	0.40
Categorical Model scale efficiency	NA	0.94 (0.08)	0.51	NA	0.94 (0.06)	0.39
Standard-CRS	5	0.75 (0.13)	0.57	4	0.78 (0.11)	0.50
Standard-VRS	18	0.79 (0.12)	0.55	17	0.82 (0.10)	0.44
Standard-scale efficiency	NA	0.96 (0.07)	0.53	NA	0.95 (0.07)	0.40
2.Academic efficiency						
Categorical Model CRS	23	0.67 (0.15)	0.75	19	0.71 (0.15)	0.75
Categorical Model VRS	68	0.76 (0.15)	0.69	55	0.77 (0.15)	0.66
Categorical Model scale efficiency	NA	0.88 (0.10)	0.63	NA	0.92 (0.09)	0.60
Standard-CRS	10	0.62 (0.15)	0.85	6	0.66 (0.14)	0.80
Standard-VRS	22	0.67 (0.15)	0.84	14	0.67 (0.15)	0.79
Standard-scale efficiency	NA	0.93 (0.04)	0.22	NA	0.99 (0.02)	0.21
3.Academic effectiveness						
Categorical Model CRS	49	0.72 (0.13)	0.75	41	0.81 (0.13)	0.62
Categorical Model VRS	84	0.75 (0.13)	0.73	82	0.85 (0.13)	0.62
Categorical Model scale efficiency	NA	0.96 (0.06)	0.75	NA	0.96 (0.05)	0.50
Standard-CRS	17	0.78 (0.13)	0.82	21	0.78 (0.13)	0.69
Standard-VRS	34	0.82 (0.14)	0.81	39	0.81 (0.13)	0.69
Standard-scale efficiency	NA	0.96 (0.04)	0.43	NA	0.97 (0.04)	0.39
4.Overall efficiency						
Categorical Model CRS	28	0.51 (0.22)	0.93	21	0.54 (0.23)	0.96
Categorical Model VRS	64	0.56 (0.24)	0.96	50	0.59 (0.23)	0.95
Categorical Model scale efficiency	NA	0.90 (0.13)	0.95	NA	0.92 (0.10)	0.90
Standard-CRS	5	0.35 (0.21)	0.99	4	0.36 (0.21)	0.98
Standard-VRS	19	0.39 (0.23)	0.99	13	0.39 (0.22)	0.98
Standard-scale efficiency	NA	0.92 (0.11)	0.81	NA	0.92 (0.09)	0.88

During both test periods, the wide range of school size in both samples (2009-2011: $range_{FS}$=2926 and $range_{PS}$=2862 with $mean_{FS}$ = 841 and $mean_{PS}$ = 990; 2004-2006: $range_{FS}$=2968 and $range_{PS}$=2917 with $mean_{FS}$ = 820 and $mean_{PS}$ = 976) results in the large variation of the indexed cost per student (2009-2011: $range_{FS}$=\$8,541 and $range_{PS}$=\$6,503 with $mean_{FS}$ = \$8,417 and $mean_{PS}$ = \$8,057; 2004-2006: $range_{FS}$=\$6,303 and $range_{PS}$=\$4,837 with $mean_{FS}$ = \$6,478 and $mean_{PS}$ = \$6,174). Comparisons of cost per student and efficiency scores by school size level for the full sample[43] in the two test periods are detailed in Appendix VI. It appears that the smallest schools (roll < 300) have the highest level of cost per student (2009-2011: mean=$11,648; 2004-2006: mean=$8,671); moreover, compared to larger schools (roll>1,200), smaller schools in each sample have much more variation in the cost per student given the significant change in standard deviation values. The variation in the cost efficiency scores is driven by school roll size and its effect on the cost per student.

Table A6.1 in Appendix VI shows that for 2009-2011, cost efficiency is positively associated with school size level; academic efficiency scores for schools with roll less than 2,400 are between 0.81 and 0.91. Academic efficiency for the largest school group (roll >2,400) is the lowest being 0.74 as a result of two group members[44] representing six school years contributing 37.5% to the group, having a small number of NCEA passes because they primarily used the CIE system; however, these schools have good reputations and a great proportion of their NCEA passes could be transferred to university entrance and higher level NCEA Level 3 endorsements (i.e. Merit and Excellence). Therefore, the small number of NCEA passes in these schools have no influence on their academic effectiveness scores. Table A6.1 demonstrates that except for the smallest schools (roll<300), school academic effectiveness is positively related to school size level. The relatively higher scores of the smallest schools are caused by sample size bias because the means of SES of the smallest schools are the lowest compared to other groups by size (see Tables 4.5 and 4.6). That is, most of them are sitting in low SES decile categories and hence their efficiency scores are calculated based on smaller sample sizes. These schools tend to derive higher scores as research reveals that average DEA score decreases with an increase in sample size (Zhang and Bartels, 1998). This explains that efficiency scores of decile one schools under the categorical model are higher than decile 2-3 schools.

---

[43] Schools in the pure secondary school sample have the same trend as the full sample.
[44] They are Auckland Grammar School and Meacleans College.

In alignment with the above analysis, except for the smallest and largest schools, the overall efficiency of schools is positively correlated to school size level. The smallest schools (roll <300) have higher efficiency scores than the second smallest schools due to sample size bias, while the largest schools (roll>2,400) are less overall efficient than the second largest schools as some use CIE as their major qualification system. Results for 2004-2006 (Table A6.2 in Appendix VI) are similar to 2009-2011 but the largest schools outperform other groups by size in terms of cost efficiency, academic effectiveness and overall efficiency, whereas academic efficiency is also at a high level. This is a consequence of most large schools adopting NCEA as the major qualification system in the earlier period.

To examine the relationship between efficiency scores and SES decile, Tables A7.1 and A7.2 in Appendix VII illustrate scores of schools in the full sample[45] under both standard and categorical VRS models in both test periods. For 2009-2011, significant positive relationships under the standard VRS model are found between the SES decile and efficiency scores for academic efficiency, academic effectiveness and overall efficiency, in keeping with the view that school academic performance positively varies with school SES decile. This is also supported by the fact that the majority of fully efficient schools are sitting in decile 9-10, particularly for academic efficiency, academic effectiveness and overall efficiency. However, the case for cost efficiency under standard VRS model is slightly different. Cost efficiency is positively related to SES decile due to the positive relationship between cost efficiency and school size, as large schools generally have higher decile. Nevertheless, the majority of efficient schools are clustered in decile four, five and six rather than decile nine or ten, given the schools with medium deciles have less opportunity to get funding from the Government or parents, compared to schools with lower or higher deciles.[46]

In contrast, efficiency under the categorical VRS model (selected as it adjusts both scale and SES effects) varies much less among deciles, the maximum range in scores for cost efficiency, academic efficiency, academic effectiveness and overall efficiency is 0.18, which is much smaller than 0.52, the upper bound of range in scores for standard VRS model. Further, more fully efficient schools are identified under the categorical VRS model compared to the standard VRS model; these schools tend to be clustered in the

[45]Similar results were received for the pure secondary school sample and CRS model.
[46] Low decile schools can get more grants from the MoE while high decile schools have more opportunities to get the donation from parents.

lowest deciles (1-3) with the remaining evenly allocated into other deciles. However, although cost efficiency and academic efficiency are even among deciles whereas academic effectiveness shows a U-shaped distribution across deciles, the overall efficiency is still positively related to decile even after student family background has been adjusted, in particular, decile ten schools have the highest overall scores. This is consistent with higher decile schools outperforming lower decile schools irrespective of whether student background factors have been adjusted.

Results for 2004-2006 are similar to those for 2009-2011. Compared to the later period, academic effectiveness scores under the categorical VRS model show a much flatter distribution among deciles. Of note, in both test periods, decile one schools outperform decile two or three schools in terms of cost efficiency, academic efficiency, academic effectiveness and overall efficiency according to the results under the categorical VRS model. As this is not in line with the view that lower decile schools underperform higher decile schools, a further review of decile one schools was undertaken. According to the results from applying the Kruskal-Wallis test, decile one schools outperform decile two and three schools in terms of NCEA passes and university entrances, whilst underperform with the ability of getting more NCEA L3 M&E. Moreover, according to the formulation of the categorical DEA model, decile one schools have the smallest sample size to calculate efficiency scores compared to others, therefore decile one schools technically have more opportunities to receive higher scores and more efficient units compared to other schools. In this sense, it is reasonable to examine the performance of decile one schools by looking at the results of both standard and categorical models; in addition, the better academic performance of decile one schools is considered as the consequence of stronger assistance and intervention from the MoE and the ERO.

4.5.1.2 Comparisons of efficiency scores by school characteristics

Prior research reveals that performance of school groups with different characteristics might differ in a specific educational context. To examine New Zealand school performance between groups with different characteristics such as school type (i.e. expanded or pure secondary schools), ownership (integrated or non-integrated schools) and gender (single sex or co-educational schools), comparisons of DEA variables and efficiency scores are made for these characteristics in Tables 4.11-4.14.

For the full sample in the period 2009-2011, expanded schools have higher mean SES decile but smaller mean roll than pure secondary schools; therefore, they have less mean teacher salaries and less mean other expenditures. However, expanded schools are less cost efficient. This is because while the roll is proportionately smaller, the inputs are reduced at a lower proportion due to the effect of the fixed costs[47] included in school operational expenditures. The importance of this difference is significant but weak ($p<0.01$; $r<0.1$)[48]. In contrast, even though the comparison for the period 2004-2006 produced similar results, the difference in cost efficiency by school type is more significant and stronger than that in the 2009-2011 sample ($p<0.001$; $0.1<r<0.2$) (see Table 4.11).

[47] For example, rates, use of lands and buildings and principal's salary etc.
[48] Significance and effect size are reported based on U-Test results.

Table 4.11 Comparison of DEA variables and categorical VRS efficiency scores by school types

Period 2009-2011						
	Expanded secondary schools (years 7-15) (n=237)		**Pure secondary schools (years 9-15) (n=628)**		**Independent *t*-test**	**Mann-Whitney U test**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**		
Decile	6.29	2.47	5.46	2.54	4.318***b	-4.196***b
Indexed teachers' salaries	2,533,835	857,061	4,628,814	2,143,468	-20.528***c	15.127***c
Indexed other expenditures	1,489,886	703,765	2,676,723	1,269,391	-17.394***c	15.068***c
Roll (year 9-15)	483	249	975	558	-17.868***c	13.46***c
Number of NCEA Level 1 passes	91	49	174	105	-15.754***c	11.937***c
Number of NCEA Level 2 passes	78	44	150	96	-14.981***c	11.205***c
Number of NCEA Level 3 passes	47	32	82	66	-10.335***c	7.522***c
Number of university entrance	43	31	74	64	-9.685***c	6.837***c
Number of NCEA Level 3 Merit and Excellence	14	15	22	29	-5.457***b	3.52***b
Cost efficiency	.7624	.1358	.7886	.1084	-2.661**b	2.846**a
Academic efficiency	.7607	.1641	.7487	.1517	1.012a	-1.269a
Academic effectiveness	.8535	.1069	.8329	.1236	2.264*a	-1.986*a
Overall efficiency	.6289	.2250	.5756	.2199	3.154**b	-2.845**a
Period 2004-2006	**Expanded schools (n=229)**		**Pure secondary schools (n=627)**			
Decile	6.29	2.65	5.58	2.55	3.555***b	-3.541***b
Indexed teachers' salaries	2,166,858	753,434	4,195,136	1,992,865	-21.606***c	15.763***c
Indexed other expenditures	722,974	333,232	1,401,914	705,173	-18.992***c	15.261***c
Roll (year 9-15)	438.38	214.08	959.52	538.52	-20.244***c	14.899***c
Number of NCEA Level 1 passes	79	42	159	102	-16.256***c	12.451***c
Number of NCEA Level 2 passes	64	39	127	94	-13.96***c	10.714***c
Number of NCEA Level 3 passes	33	26	63	59	-10.433***c	7.833***c
Number of university entrance	31	25	59	56	-9.919***c	7.484***c
Number of NCEA Level 3 Merit and Excellence	8	10	16	22	-6.939***c	5.178***b
Cost efficiency	.7916	.1292	.8341	.1013	-4.501***c	4.163***b
Academic efficiency	.8046	.1391	.8204	.1265	-1.581a	1.485a
Academic effectiveness	.7770	.1464	.7578	.1512	1.66a	-1.661a
Overall efficiency	.5870	.2600	.5554	.2327	1.621a	-1.52a

Significance: ****p* <0.001, ***p*<0.01 (2- tailed); effect size[49] (*r*), a=0.01<*r*<0.1, b=0.1<*r*<0.2, c≥0.2

[49] Effect size is a measure of how important that the significant differences between two practical terms are (Field, 2013, p.376). Effect size for *t*-test, $r=\sqrt{\frac{t^2}{t^2+df}}$; Effect size for U-test is denoted by r, $r=\frac{z}{\sqrt{N}}$.

Given that the results for the pure secondary school sample are similar to that for the full sample, Tables 4.12-4.14 only present the statistics for the comparisons of DEA results between groups by school characteristics for the full sample.

Table 4.12 shows that for both 2009-2011 and 2004-2006, integrated schools have higher mean SES decile but smaller mean roll than non-integrated schools; therefore, they have less mean teacher salaries and less mean other expenditures. Similar to the expanded schools, integrated schools are less cost efficient than non-integrated schools due to the effect of the fixed costs included in school operational expenditures. The importance of this difference is significant but relatively weak ($p<0.01$; $0.1<r<0.2$). However, integrated schools are more academic efficient, academic effective and hence overall efficient compared to non-integrated schools. This is in keeping with the findings of prior studies that integrated schools still preserve an edge over many non-integrated schools in terms of effective management practices. This may be partially related to the strong support from their community of interest stemmed from the specific ethos or beliefs of such schools (Alexander et al., 2010).

Table 4.12 Full sample: comparison of DEA variables and efficiency scores by school ownership

Period 2009-2011						
	Integrated schools (n=174)		**Non-integrated schools (n=691)**		**Independent *t*-test**	**Mann-Whitney U test**
	Mean	**Std. Dev**	**Mean**	**Std. Dev**		
Decile	6.86	2.48	5.39	2.48	6.988***c	-6.987***c
Indexed teachers' salaries	2,672,407	776,611	4,402,917	2,183,158	-16.999***c	10.624***c
Indexed other expenditures	1,604,178	717,028	2,539,736	1,296,018	-12.748***c	10.061***c
Roll (year 9-15)	513	202	923	566	-15.497***c	9.232***c
Number of NCEA Level 1 passes	101	41	164	106	-12.352***c	6.887***c
Number of NCEA Level 2 passes	89	38	141	97	-10.904***c	5.669***b
Number of NCEA Level 3 passes	59	30	76	66	-4.984***b	0.77
Number of university entrance	53	28	69	64	-4.912***b	0.354
Number of NCEA Level 3 Merit and Excellence	19	16	20	28	-0.856	-2.883**a
Cost efficiency	.7533	.1313	.7885	.1122	-3.243*b	3.513***b
Academic efficiency	.7902	.1423	.7423	.1569	3.665***b	-3.56***b
Academic effectiveness	.8941	.0965	.8246	.1207	8.051***c	-7.585***c
Overall efficiency	.6995	.1997	.5627	.2196	7.478***c	-7.193***c
Period 2004-2006	**Integrated schools (n=183)**		**Non-integrated schools (n=673)**			
Decile	6.63	2.72	5.54	2.51	5.12***b	-5.153***b
Indexed teachers' salaries	2,189,768	658,216	4,050,272	2,015,110	-20.298***c	13.009***c
Indexed other expenditures	724,809	276,705	1,355,009	714,280	-18.373***c	12.684***c
Roll (year 9-15)	456	190	919	546	-18.299***c	11.699***c
Number of NCEA Level 1 passes	88	40	151	103	-12.719***c	7.799***c
Number of NCEA Level 2 passes	74	37	120	94	-10.037***c	5.422***c
Number of NCEA Level 3 passes	40	24	59	59	-6.721***c	2.341*a
Number of university entrance	38	24	55	56	-6.347***c	2.144*a
Number of NCEA Level 3 Merit and Excellence	10	10	15	22	-4.559***b	1.292a
Cost efficiency	.8096	.1200	.8263	.1083	-1.815*a	1.376a
Academic efficiency	.8238	.1220	.8141	.1322	0.89a	-0.69a
Academic effectiveness	.8486	.1287	.7396	.1470	9.849***c	-9.227***c
Overall efficiency	.6718	.2347	.5345	.2338	7.036***c	-6.868***c

Significance: ****p* <0.001, ***p*<0.01 (2- tailed); effect size[50] (r), a=0.01<*r*<0.1, b=0.1<*r*<0.2, c≥0.2

[50] Ibid.

Table 4.13 illustrates that for both 2009-2011 and 2004-2006, single-sex schools have higher mean SES decile and larger mean roll than co-educational schools. However, the differences in mean teachers' salaries and mean other operational expenditures are not significant. It appears that single-sex schools are more academic efficient, academic effective and overall efficient than co-educational schools. This finding is in alignment with the prior research, given that the curriculum settings and teaching practices in single-sex schools cater only for boys or girls and hence are more effective than their co-educated counterparts. However, results for 2004-2006 show that single-sex schools are more cost efficient than their co-educated peers whereas this feature is not significant for 2009-2011. The possible reason is that in the earlier period, teachers' salaries and other expenditures for single-sex schools are on average less than co-educational schools while in the later period, the total expenditures are higher (Table 4.13).

Table 4.13 Full sample: comparison of DEA variables and efficiency scores by single sex/co-educational schools

Period 2009-2011						
	Single-sex schools (n=257)		Co-educational schools (n=608)		Independent T-test	Mann-Whitney U test
	Mean	Std. Dev	Mean	Std. Dev		
Decile	6.90	2.31	5.18	2.47	9.816***c	-9.173***c
Indexed teachers' salaries	4,082,185	1,843,519	4,043,245	2,200,402	0.268a	-1.719a
Indexed other expenditures	2,419,308	1,271,482	2,322,899	1,253,678	1.029a	-1.203a
Roll (year 9-15)	872	481	827	563	1.199a	-2.824**a
Number of NCEA Level 1 passes	163	86	147	105	2.421*a	-4.348***b
Number of NCEA Level 2 passes	144	79	124	95	3.134**b	-5.324***b
Number of NCEA Level 3 passes	89	56	66	62	5.175***b	-8.136***c
Number of university entrance	82	55	58	59	5.663***b	-8.829***c
Number of NCEA Level 3 Merit and Excellence	27	26	18	25	5.556***b	-9.226***c
Cost efficiency	0.7857	0.1151	0.7796	0.1179	0.702a	-1.429a
Academic efficiency	0.7823	0.1344	0.7391	0.1616	4.051***b	-3.546***b
Academic effectiveness	0.8806	0.1101	0.8208	0.1189	7.139***c	-7.932***c
Overall efficiency	0.6879	0.1990	0.5489	0.2191	8.753***c	-8.485***c
Period 2004-2006	**Single-sex schools (n=268)**		**Co-educational schools (n=588)**			
Decile	6.81	2.532	5.3	2.486	8.214***c	-7.809***c
Indexed teachers' salaries	3,527,741	1,706,494	3,709,398	2,072,797	-1.348a	0.34a
Indexed other expenditures	1,186,589	652,412	1,235,638	714,702	-0.956a	0.629a
Roll (year 9-15)	818	471	821	551	-0.067a	-1.075a
Number of NCEA Level 1 passes	147	83	133	102	2.082*a	-4.091***b
Number of NCEA Level 2 passes	125	78	104	91	3.277**b	-5.777***b
Number of NCEA Level 3 passes	68	50	50	55	4.555***b	-7.742***c
Number of university entrance	64	47	50	52	4.902***b	-8.409***c
Number of NCEA Level 3 Merit and Excellence	18	20	12	20	4.16***b	-7.755***c
Cost efficiency	0.8375	0.1064	0.8161	0.1125	2.681***b	-3.279**b
Academic efficiency	0.8599	0.0935	0.7962	0.1393	7.862***c	-6.48***c
Academic effectiveness	0.827	0.1340	0.7337	0.1480	9.145***c	-8.843***c
Overall efficiency	0.6836	0.2077	0.5092	0.2348	10.926***c	-10.159***c

Significance: ****p* <0.001, ***p*<0.01 (2- tailed); effect size[51] (*r*), a=0.01<*r*<0.1, b=0.1<*r*<0.2, c≥0.2

[51] Ibid.

To further examine the differences in performance between single-sex and co-educational schools, both parametric one-way ANOVA and non-parametric Jonckheere-Terpstra[52] were employed to look at differences between three groups: boys, girls and co-educational schools (Table 4.14). Kruskal-Wallis test provided similar results to one-way ANOVA but the effect size for testing areas are hard to compute (Field, 2013). In this circumstance, Jonckheere-Terpstra is applied to look at the importance of the differences between groups.

Table 4.14 demonstrates that in both test periods, both boys and girl schools have higher mean SES decile than their co-educated peers. Girl schools outperform both boy schools and co-educational schools in terms of academic efficiency, academic effectiveness and overall efficiency, which is consistent with the finding of prior research that girls are more capable of getting higher academic achievements (Alexander et al., 2010). With regards to the differences between the two test periods, on one hand, boy schools are more cost efficient than co-educational schools during 2004-2006 whereas this is not significant during 2009-2011; on the other hand, in the earlier test period, boy schools outperform co-educational schools in terms of academic efficiency and overall efficiency whereas this is not significant in the later period. The differences relating to boy school performance between the two test periods indicate that the NCEA system may be reducing inequity across different groups of students, or boys perform less well under the new system.

[52] The Jonckheere–Terpstra test is a test for an ordered alternative hypothesis regarding independent groups within a sample. The main difference between the Jonckheere–Terpstra test and the Kruskal–Wallis test is that, the Jonckheere–Terpstra test has a priori ordering of independent groups whereas the Kruskal-Wallis test has not. In this study, there is no priori ordering required and Jonckheere–Terpstra test was employed to estimate the effect size while comparing the differences between groups, given it produced similar results to the Kruskal-Wallis test which does not provide *z*-value for the calculation of effect size.

Table 4.14 Full sample: comparison of DEA variables and efficiency scores by boys/girls/co-educational schools

Period 2009-2011	**Boy Schools (n=117); Girl Schools (n=140); Co-educational Schools (n=608)**			
Areas with significant difference in mean	**Groups compared**	**Mean difference**	**One-way ANOVA[53]**	**Jonckheere-Terpstra Test[54]**
SES decile	Boy > Co-ed Girl > Co-ed	1.80 1.66	45.68***c	-9.10***c
Number of NCEA Level 3 passes	Girl > Boy Girl > Co-ed	20 32	17.19***b	-7.52***c
Number of university entrance	Boy > Co-ed Girl > Boy Girl > Co-ed	15 18 32	19.10***b	-8.28***c
Number of NCEA Level 3 Merit and Excellence	Girl > Boys Girl > Co-ed	3 2	24.06***c	-8.35***c
Academic efficiency	Girl > Boy Girl > Co-ed	0.107 0.092	23.33***c	-2.37*a
Academic effectiveness	Girl > Boy Girl > Co-ed	0.087 0.100	43.65***c	-6.66***c
Overall efficiency	Girl > Boy Girl > Co-ed	0.232 0.245	83.16***c	-6.71***c
Period 2004-2006	**boy schools (n=122); girl schools (n=146); co-educational schools (n=588)**			
SES decile	Boy > Co-ed Girl > Co-ed	0.249 0.231	33.8***c	-7.59***c
Number of NCEA Level 2 passes	Girl > Co-ed	23.72	5.521**b	-5.6***b
Number of NCEA Level 3 passes	Girl > Boy Girl > Co-ed	17 26	13.87***b	-7.1***c
Number of university entrance	Girl > Co-ed	25	14.53***b	-7.84c
Number of NCEA Level 3 Merit and Excellence	Girl > Boy Girl > Co-ed	2.4 1.8	14.5***b	7.0***c
Cost efficiency	Boy > Co-ed	0.029	3.97*a	-3.38**b
Academic efficiency	Boy > Co-ed Girl > Boy Girl > Co-ed	0.05 0.024 0.074	24.5***c	-6.01***c
Academic effectiveness	Girl > Boy Girl > Co-ed	0.112 0.144	61.88***c	-7.46***c
Overall efficiency	Boy > Co-ed Girl > Boy Girl > Co-ed	0.06 0.209 0.27	88.67***c	-8.47***c

Significance: ****p* <0.001, ***p*<0.01 (2- tailed); effect size[55] (*r*), a=0.01<*r*<0.1, b=0.1<*r*<0.2, c≥0.2

[53] This column records the *F*-values. Effect size= $\sqrt{\frac{sum\ of\ squares\ between\ groups}{total\ sum\ of\ squares}}$.

[54] This column reports *z*-values. Effect size= $\frac{z}{\sqrt{N}}$.

[55] Ibid.

4.5.1.3 Discussion on efficient schools

Further to the discussion in the last section, Tables A8.1-A8.4 in Appendix VIII demonstrate the features of efficient schools by characteristic. Categorical VRS results are selected as they have removed the effects of both SES and scale.

For 2009-2011, Table A8.1 illustrates that most efficient schools are low decile (1-3) or decile ten schools. It is likely that decile ten schools outperform schools in other deciles because school academic achievement is positively associated with SES decile. As expected, there are more efficient schools sitting in decile ten. However, more efficient schools are clustered in the lowest deciles under the categorical model, as the scores for the lowest deciles are generated based on the smallest sample sizes during the nesting process of calculation. With the same inputs and outputs, a smaller sample tends to have higher percentage of efficient DMUs than a larger one.

Table A8.1 also displays the numbers of multiple efficient schools in terms of cost efficient and academic efficient, cost efficient and academic effective, academic efficient and effective and all efficient with the above three efficiencies. It appears that the roll variation of multiple efficient schools is large and a further review of these schools indicated that larger schools tend to be more cost efficient and academic efficient whereas smaller schools tend to be more academic effective and overall efficient as school roll is not directly included as a DEA variable for the calculation of academic effectiveness and overall efficiency. As such these characteristics are potentially reflected via the large standard deviations of scores for the multiple efficient schools.

Statistics of efficient schools by characteristics show the following findings:

- Larger proportion[56] of expanded schools are efficient compared to pure secondary schools;
- Larger proportion[57] of integrated schools are efficient than non-integrated schools;
- Larger proportion[58] of girl schools are academic efficient, academic effective, overall efficient and both academic efficient and effective, relative to their portion in the sample;

[56] The percentage of expanded schools in efficient schools is larger than the percentage in the sample (27.4%).
[57] The percentage of integrated schools in efficient schools is larger than the percentage in the sample (20.1%).
[58] The percentage of girl schools in efficient schools is larger than the percentage in the sample (16%).

- Smaller proportion[59] of boy schools are efficient for all the efficiency types, relative to their portion in the sample;
- Larger proportion[60] of co-educational schools are efficient except for overall efficiency, relative to their portion in the sample.

Statistics for pure secondary school sample in 2009-2011 (see Table A8.2) show similar features to the full sample except for the differences as follows: first, there are no expanded schools in the pure secondary schools sample therefore the number of efficient schools for expanded schools is not applicable; second, larger proportion[61] of girl schools are efficient for all efficiency types and finally, larger proportion of co-educational schools are cost efficient or both cost efficient and academic efficient.

Results for 2004-2006 are similar to that for 2009-2011 (see Tables A8.3-A8.4), but the superiority of girl schools is weaker compared to the later test period.

Tables A9.1-A9.4 in Appendix IX provide the information (categorical VRS) to examine whether the relationships among variables for efficient schools in one sample are different from those in the whole sample. Given that the pure secondary school sample produces similar results, these tables only focus on the full sample.

Statistics show that cost efficient schools are larger (2009-2011: mean roll=964; 2004-2006: mean roll=1,133) with lower cost per student in each corresponding decile. Academic efficient schools are larger (2009-2011: mean roll=1,067; 2004-2006: mean roll=999) and more efficient at obtaining higher NCEA Level 1-3 qualification pass rates. The possible reason for this is that large schools, though some are in low decile, generally have good reputation for school academic achievement and thus attract more students to enrol. Results for academic effective schools in 2009-2011 have the same level mean roll as the entire sample whereas higher in 2004-2006 (2009-2011: mean roll=881; 2004-2006: mean roll=944). However, academic effective schools have significantly higher pass rates for UE and NCEA L3 M&E, which are considered to be essential for university achievement. Likewise, compared to the entire sample, overall efficient schools are not significantly larger (2009-2011: mean roll=841; 2004-2006:

[59] The percentage of boy schools in efficient schools is smaller than the percentage in the sample (13.5%).
[60] The percentage of co-educational schools in efficient schools is larger than the percentage in the sample (70.3%).
[61] The percentage of girl schools in the pure secondary school sample is 16.6% while the percentage of co-educational schools in the sample is 70.4%.

mean roll=952), but are more academic effective in terms of achieving much higher rates for UE passes and NCEA L3 M&E endorsements.

It is noteworthy that the overall efficient schools in 2009-2011 have higher mean indexed cost per student compared to schools in the full sample. Further examination revealed that compared to schools in the full sample, the overall efficient schools have increases in the pass rates of UE and NCEA L3 M&E endorsements, which are greater than the increase in cost per students. As a result, the overall efficient schools are more capable of generating academic outcomes from limited school resources, though sometimes consuming more expenditures than the average level.

4.5.1.4 Comparison between overall efficiency and sub-efficiencies

As mentioned in section 2.4.4, a network DEA model is employed in this study to measure school performance from multiple dimensions by including cost efficiency, academic efficiency and academic effectiveness sub-models into the performance measurement process. The primary advantage of applying a DEA network model in school performance measurement system is that, it can identify the sources of inefficiency in parts of the educational process. This can be confirmed by contrasting the sub-efficiencies (i.e., cost efficiency, academic efficiency and academic effectiveness) with the overall efficiency, which is considered a bridged model that incorporates the original inputs and the final outputs of the process. In this regard, this section first examines the sub-efficiencies for efficient schools identified by the overall efficiency model, then looks at whether the bottom 10%[62] inefficient schools identified by the overall efficiency model are also classified as being the lowest 10% inefficient by the sub-models.

Table 4.15 lists the numbers of schools with 100% sub-efficiency for all the efficient schools with 100% overall efficiency. For both the test periods and for both the full sample and the pure secondary school samples, all the mean efficiencies are high being over 0.9 with standard deviations within 0.13. Efficiencies for the pure secondary schools are higher with lower standard deviations compared to schools in the full sample, given the pure secondary school sample is more homogeneous. According to the numbers of schools with 100% sub-efficiency as well as the sub-efficiency scores,

[62] 10% is selected to cover inefficient schools for an approximately equivalent number of efficient schools. For example, for the full sample in 2009-2011, the number of bottom 10% overall inefficient schools is equal to 87, given there are 70 overall efficient schools.

the overall efficient schools tend to be more academic efficient and academic effective. However, only a small proportion of overall efficient schools are all efficient in sub-efficiencies whereas some overall efficient schools are not efficient in any of the sub-efficiencies. For example, for the full sample during 2009-2011, only 10% of the overall efficient schools are efficient in all the sub-efficiencies, whereas 15.7% of overall efficient schools are not efficient in any of the cost efficiency, academic efficiency and academic effectiveness sub-models.

It is interesting that two school years in each test period are fully efficient for all the sub-efficiencies but not overall efficient (see remarks in Table 4.15). This is a rare case for network DEA model which is designed to look at efficiency of DMUs via a pipeline consisting of seamlessly connected sub-efficiency models. However, for this study, the overall efficiency model ‘bridges over’ other performance measures which reflect school performance from different perspectives, such as the numbers of NCEA passes. This would result in some overall efficient schools (not all efficient in sub-efficiencies) better utilizing school resources to achieve higher levels of outcomes compared to a few schools that are fully efficient for all the sub-efficiencies. Table 4.15 shows that school year 20090088, 20100099, 20040028 and 20040074 are all efficient in the sub-efficiencies but not overall efficient. A review of their 100% overall efficient peers showed that some peers are not 100% cost efficient but scored more than 88%; however, all these peers are 100% academic effective with higher levels of outputs relative to the same level of the inputs, compared to the above four school years with 100% sub-efficiencies.

Table 4.15 Sub-efficiency status for overall efficient schools (categorical VRS)

Test period 2009-2011						
	Full sample overall efficient schools (n=70)			**Pure secondary school sample overall efficient schools (n=49)**		
	Cost efficiency	**Academic efficiency**	**Academic effectiveness**	**Cost efficiency**	**Academic efficiency**	**Academic effectiveness**
Efficiency Mean (St. Dev)	0.9021 (0.110)	0.9391 (0.091)	0.9413 (0.091)	0.9170 (0.091)	0.9694 (0.071)	0.9551 (0.074)
# Efficient schools	22	34	41	15	33	31
# Multiple efficient schools	CEc&Aec: 8			10		
	Cec&Aet: 11			10		
	Aec&Aet: 26			22		
	Cec, Aec&Aet: 7			8		
Remarks	20090088 and 20100099 are all sub-efficient but not overall efficient[63]; some overall efficient schools are not efficient in any sub-efficiencies (FS:15.7%; PS: 8.2%).					
Test period 2004-2006						
	Full sample Overall efficient schools (n=64)			**Pure secondary school sample Overall efficient schools (n=50)**		
	Cost efficiency	**Academic efficiency**	**Academic effectiveness**	**Cost efficiency**	**Academic efficiency**	**Academic effectiveness**
Efficiency Mean (St. Dev)	0.9217 (0.097)	0.9261 (0.121)	0.9475 (0.093)	0.9219 (0.087)	0.9587 (0.079)	0.9710 (0.063)
# Efficient schools	22	32	37	15	32	33
# Multiple efficient schools	Cec&Aec: 10			8		
	Cec&Aet: 12			11		
	Aec&Aet: 23			24		
	Cec, Aec&Aet: 7			6		
Remarks	20040028 and 20040074 are all sub-efficient but not overall efficient[64]; some overall efficient schools are not efficient in any sub-efficiencies (FS:17.2%; PS: 14%)					

[63] These school years have the same features for both the samples.
[64] Ibid.

In contrast, Table 4.16 considers the status of sub-efficiencies for the bottom 10% inefficient schools identified by the overall efficiency model. During the period 2009-2011, for both samples, numbers of the bottom 10% inefficient schools are even in terms of cost efficiency, academic efficiency and academic effectiveness; whereas during 2004-2006, for both samples, a large proportion of the bottom 10% overall inefficient schools are also bottom 10% academic ineffective, followed by a lower proportion sitting in the bottom 10% academic inefficient group. This indicates an increase in school academic performance from the earlier period to the later period. In addition, Table 4.16 also shows that only a small proportion of bottom 10% overall inefficient schools are all bottom 10% inefficient in terms of cost efficiency, academic efficiency and academic effectiveness. For example, for the full sample during 2009-2011, only 1.1% bottom 10% overall inefficient schools are all bottom 10% inefficient in sub-efficiencies, whereas 30% of the bottom 10% overall inefficient schools are not bottom 10% inefficient in any of the sub-efficiencies.

Table 4.16 Sub-efficiency status for the bottom 10% inefficient schools (categorical VRS)

Test period 2009-2011						
	Full sample Bottom 10% overall inefficient schools (n=87)			**Pure secondary school sample Bottom 10% overall inefficient schools (n=62)**		
	Cost efficiency	**Academic efficiency**	**Academic effectiveness**	**Cost efficiency**	**Academic efficiency**	**Academic effectiveness**
# Bottom 10% inefficient schools	21	30	26	16	24	18
# Multiple inefficient schools (bottom 10%)	Cec&Aec: 3			2		
	Cec&Aet: 7			5		
	Aec&Aet: 7			6		
	Cec, Aec&Aet: 1			1		
Remarks	Some bottom 10% overall inefficient schools are not bottom 10% inefficient with any of the sub-efficiencies (FS: 30%; PS: 26%)					
Test period 2004-2006						
	Full sample Bottom 10% overall inefficient schools (n=86)			**Pure secondary school sample Bottom 10% overall inefficient schools (n=61)**		
	Cost efficiency	**Academic efficiency**	**Academic effectiveness**	**Cost efficiency**	**Academic efficiency**	**Academic effectiveness**
# Bottom 10% inefficient schools	14	33	53	11	25	41
# Multiple inefficient schools (bottom 10%)	CEc&Aec: 4			2		
	Cec&Aet: 9			6		
	Aec&Aet: 20			16		
	Cec, Aec&Aet: 3			1		
Remarks	Some bottom 10% overall inefficient schools are not bottom 10% inefficient with any of the sub-efficiencies (FS: 17%; PS: 11%)					

It appears that as a single standard DEA model, overall efficiency model treats the process of school performance measurement as a 'black box', which hides some critical information for performance evaluation and improvement, both for schools classified as efficient and inefficient. For each individual school under observation, the overall efficiency model fails to demonstrate how the sub-processes in each part of the educational process contribute to its performance evaluation. For the overall efficient schools, it hides some areas for improvement to which these schools should pay attention; for inefficient schools, it fails to recognize areas of good or poor practice for these schools to share good experience in the educational system or improve their performance with the supports from their efficient peers. In this regard, this study employed three sub-efficiency models to decompose the whole performance measurement process, in order to capture more performance information from multiple facets for each school under observation and promote the learning network.

4.5.1.5 Comparison of efficiencies between years for each test period

To examine the differences in efficiencies for each test period, both one-way ANOVA and the Kruskall-Wallis test were employed (not reported here). For 2009-2011, results from ANOVA show there were no significant differences in any types of efficiency whilst those from Kruskall-Wallis test indicate that there was significant increase ($p<0.05$) in school academic efficiency and effectiveness in the later years of this period (i.e., 2010 and 2011).

In contrast, for 2004-2006, results from both ANOVA and the Kruskall-Wallis test indicate that school cost efficiency significantly declined over the years ($p<0.001$), as a consequence of the increased government grants that helped schools with software for managing NCEA data, NCEA related operational costs and NCEA professional development for teachers (NZ Herald, 2003). Moreover, academic efficiency in 2005 is on average lower than that in 2004; however, in 2006, academic efficiency increased significantly ($p<0.001$), better than both 2004 and 2005. Nevertheless, differences in academic effectiveness and overall efficiency between the years in this period were not significant.

4.5.2 Returns to scale

Given the wide range of school size existing in both samples, issues relating to economies of scale might arise and therefore returns to scale were examined. Tables 4.17-4.18 display the details for the returns to scale (RTS) under the categorical VRS model. The number of schools with increasing returns to scale (IRS), decreasing returns to scale (DRS) and constant returns to scale (CRS) together with the corresponding mean roll for each category are provided.

As shown in Tables 4.7 and 4.8, all the inputs and outputs in the network DEA model are significantly related to school roll, as such it is reasonable to examine returns to scale in conjunction with roll size. Tables 4.17 and 4.18 illustrate that small schools have IRS whilst large schools have CRS. There are no DRS identified in both samples and this implies that New Zealand secondary school performance could be improved via the increase in scale measured by school size/roll.

Statistics for both test periods indicate that the primary difference in RTS results between the full sample and the pure secondary school sample is that the former changes from IRS to CRS at a smaller increase in roll size compared to the latter. This is the consequence that the full school sample contains many expanded schools (i.e. years 7-15 schools) that are normally small, whereas the pure secondary school sample includes much more schools with large rolls.

Table 4.17 Returns to scale for categorical VRS model period 2009-2011

	Cost efficiency model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	818 94.6%	822.34	504.0	577 93.8%	962.28	513.2
CRS	47 5.4%	1157.1	924.5	38 6.2%	1418.4	894.1
DRS	0	0	0	0	0	0
Total	865	840.53	539.9	615	990.47	554.4

	Academic efficiency model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	779 90%	809.49	479.6	544 88.5%	949.57	488.0
CRS	86 10%	1121.7	876.7	71 11.5%	1303.8	858.0
DRS	0	0	0	0	0	0
Total	865	840.53	539.9	615	990.47	554.4

	Academic effectiveness model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	771 89.1%	840.51	498.3	537 87.3%	988.6	503.6
CRS	94 10.9%	840.71	807.7	78 12.7%	1003.3	827.9
DRS	0	0	0	0	0	0
Total	865	840.53	539.9	615	990.47	554.4

	Overall efficiency model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	779 90%	829.9	506.4	550 89.4%	967.8	515.5
CRS	86 10%	936.9	778.1	65 10.6%	1182.1	791.9
DRS	0	0	0	0	0	0
Total	865	840.53	539.9	615	990.47	554.4

Table 4.18 Returns to scale for categorical VRS model period 2004-2006

	Cost efficiency model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	797 93.1%	795.1	485.3	568 93%	939.45	486.5
CRS	59 6.9%	1157.4	859.6	43 7%	1452.7	824.9
DRS	0	0	0	0	0	0
Total	856	820.1	527.1	611	975.58	533.0

	Academic efficiency model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	786 91.8%	801.56	482.5	554 90.7%	952.18	480.8
CRS	70 8.2%	1028.3	863.8	57 9.3%	1203.0	868.3
DRS	0	0	0	0	0	0
Total	856	820.1	527.1	611	975.58	533.0

	Academic effectiveness model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	772 90.2%	804.8	482.6	527 86.2%	968.74	478.4
CRS	84 9.8%	960.6	822.6	84 13.8%	1018.4	797.4
DRS	0	0	0	0	0	0
Total	856	820.1	527.1	611	975.58	533.0

	Overall efficiency model					
	Full sample			Pure secondary school sample		
	# School years (%)	Roll mean	Roll Std Dev.	# School years (%)	Roll mean	Roll Std Dev.
IRS	775 90.5%	803.17	496.8	544 89%	957.33	498.7
CRS	81 9.5%	982.11	742.8	67 11%	1123.7	744.9
DRS	0	0	0	0	0	0
Total	856	820.1	527.1	611	975.58	533.0

The most productive scale size appears where scale efficiency score is 100% using both CRS and VRS models. Table 4.19 details the statistics of roll sizes for schools operating at the most productive scale in terms of all the efficiencies. Results show that for both samples, only very small schools can operate at 100% productive scale for academic efficiency. This conflicts against the view that larger schools tend to be more academic efficient because large schools normally have high decile. The unexpected result is caused by the inclusion of roll as an output for the cost efficiency model whereas taking it as the only input for the academic efficiency model. As a result, the increase in roll leads to an increase in cost efficiency but a decrease in academic efficiency. Therefore, only the small schools can get full scale efficiency to be fully academic efficient.

To reduce the distortion relating to productive scale caused by the sole intermediary variable (i.e., roll) that connects two elements in the pipe line of the productive process, efficiency scores for the ability of transferring school resources to academic outputs[65] were calculated to examine the operating scale that can maximize academic outputs (i.e., NCEA Level 1-3 qualification passes) from the budgeted school resources. This is sensible as in practice, operating scale relates to the origin of the whole productive process rather than the transitional parts. Accordingly, the operating scale for schools to achieve best academic outputs or outcomes should be investigated by modelling the efficiency of generating the academic outputs or outcomes from the operational expenditures, given that the productive process begins with school resources. Table 4.20 shows that for both test periods, the pure secondary school sample has a larger size of the most productive scale than the full sample. When school rolls are over 900 students, on the whole, the most productive scale size occurs for producing all the academic outputs and outcomes.

[65] Efficiency scores were calculated using categorical VRS by including teachers' salaries and other resources as inputs, with numbers of NCEA Level 1-3 passes as outputs.

Table 4.19 Roll sizes for scale efficient schools by efficiency models

	Test period: 2009-2011										Test period: 2004-2006									
	Full sample (n=865)					Pure secondary school sample (n=615)					Full sample (n=856)					Pure secondary school sample (n=611)				
	N[66]	Min	Max	Mean	St. Dev.	N	Min	Max	Mean	St. Dev.	N	Min	Max	Mean	St. Dev.	N	Min	Max	Mean	St. Dev.
Cost efficient	14	404	2,688	1,397	729	15	657	2,688	1,587	633	19	471	2,242	1,176	516	17	497	2,242	1,302	454
Academic efficient	19	148	746	381	190	22	184	2,232	686	526	23	129	2,324	411	458	19	202	2,324	704	699
Academic effective	49	120	2,688	596	550	39	266	2,688	925	664	50	130	2,466	667	523	42	202	2,490	888	652
Overall efficient	25	306	2,233	942	591	28	273	2,596	1,391	688	28	139	2,591	998	773	21	202	2,605	1,446	827
Sample roll	865	120	3,046	841	540	615	184	3,046	990	554	856	129	3,097	820	527	611	180	3,097	976	533

Table 4.20 Roll sizes for scale efficient schools based on costs

	Test period: 2009-2011										Test period: 2004-2006									
	Full sample (n=865)					Pure secondary school sample (n=615)					Full sample (n=856)					Pure secondary school sample (n=611)				
	N	Min	Max	Mean	St. Dev.	N	Min	Max	Mean	St. Dev.	N	Min	Max	Mean	St. Dev.	N	Min	Max	Mean	St. Dev.
Costs to Student number	14	404	2,688	1,397	729	15	657	2,688	1,587	633	19	471	2,242	1,176	516	17	497	2,242	1,302	454
Costs to academic outputs	37	404	2,232	1,098	560	35	273	2,596	1,435	589	57	202	2,605	981	678	52	202	2,605	1,200	674
Costs to academic outcomes	25	306	2,233	942	591	28	273	2,596	1,391	688	28	139	2,591	998	773	21	202	2,605	1,446	827
Sample roll	865	120	3,046	841	540	615	184	3,046	990	554	856	129	3,097	820	527	611	180	3,097	976	533

[66] The uppercase N in Tables 4.19-4.20 refers to the number of 100% scale-efficient schools.

4.5.3 Non-radial slacks and variable weights

Non-radial slack is defined as the distance from the point where an inefficient DMU is projected onto the frontier, to the closest efficient DMU. A large amount of non-radial slacks suggests that a frontier does not envelop data well as the consequence of insufficiencies in sample sizes (Coelli, Rao, O'Donnell, & Battese, 2005).

Tables A10.1 and A10.2 in Appendix X show the non-radial slacks on the inputs and outputs of cost efficiency, academic efficiency, academic effectiveness and overall efficiency models applying categorical VRS approach. Given that the DEA models are output-oriented, non-radial slacks are only examined for the output variables. As such, for both test periods, lower levels of slack were found for NCEA Level 1 passes and UE passes, whilst higher levels of slack were obtained for the other output variables, with the highest level related to the number of NCEA L 3 M&E. This is in line with the fact that a high level of variability exists for higher level academic outputs[67]/outcomes[68] among schools. In particular, there is a wide range in the pass rate of NCEA L3 M&E endorsements for schools in both samples for both test periods, given a number of schools do not perform well with NCEA L3 M&E.

The MoE has not provided a guideline with respect to which academic outputs/outcomes are more important, nor has the ERO published performance indicators for school academic achievements. Therefore, weight restrictions to inputs/outputs were not applied to any models in this study. Accordingly, efficiency scores for each school are able to reflect the school's capability of doing what it does the best even though it may have different objectives or focuses from other schools.

As a result of not applying weight restrictions, some efficiency scores for a specific school may be derived from zero-weighted inputs/outputs. As such, the combination of input and output weights for an individual school contributes to the best efficiency score for this school. Tables A11.1 and A11.2 in Appendix XI display the weights for all the inputs and outputs in the network DEA model using the categorical VRS mode. Given the output-orientated nature of the DEA model, the following discussion only focuses on the non-zero weights for output variables.

---

[67] Higher level academic outputs refer to NCEA qualifications above Level 2.
[68] Higher level academic outcomes specifically refer to Merit and Excellence endorsements of NCEA Level 3.

For both samples during both test periods, several findings are outlined as follows:

- In the academic efficiency model, the number of non-zero weights for NCEA Level 1 passes is at high level (>84%), with lower level (52-70%) for NCEA Level 2 and Level 3 passes, indicating that most schools are good at achieving NCEA Level 1 qualification while performing less well in NCEA Level 2 and Level 3 achievements;
- In the academic effectiveness model, the number of non-zero weights for UE passes is at high level (>95%), with a lower level (73-75%) for the number of NCEA L3 M&E endorsements, in keeping with that the wide variety in the numbers of NCEA L3 M&E exists among schools and the number of UE contributes the major part of the outputs to academic effectiveness;
- In the overall efficiency model, the number of non-zero weights for UE passes is at high level (>85%), with a much lower level for the number of NCEA L3 M&E endorsements (21-32%). This demonstrates that most schools perform well in terms of maximizing UE passes while controlling the consumption of school expenditures, but the number of NCEA L3 M&E endorsements is weighted significantly less where many schools perform poorly.

4.5.4 Changes in school productivity since the introduction of NCEA

In contrast with section 4.5.1.5 which examines efficiency change across years in each test period, this section concerns the second research question, that is, whether there are any changes in school productivity since the introduction of NCEA in 2004[69]; if so, what the characteristics for these changes are. To answer this question, Malmquist indices were calculated to determine whether changes in efficiencies exist from 2004 to 2011 and if any significant changes are identified, what causes these changes, i.e., change in relative efficiency or productive frontier shift.

Since section 2.4.3 has discussed the formulation of Malmquist index, Table 4.21 displays the key components of Malmquist index.

[69] NCEA Level 1 was introduced in 2002 and during 2003-2004, NCEA Level 2 and 3 were introduced. By the end of 2004, the implementation of the whole system was completed (New Zealand Qualifications Authority, 2015). As such, this study takes 2004 as the first year for the operation of NCEA in New Zealand education system.

Table 4.21 Summary of Malmquist formulation

Output distance function	$D_0^t(y^t, x^t) = \min\{\boldsymbol{\delta}: (x^t, \frac{y^t}{\delta}) \in S^t\}$ Where, for time period *t*, S^t consists of all the combinations of input (x) and output (y) being feasible in the production process. $D_0^t(y^t, x^t)$ is generally reflected by the efficiency at time *t*.		
Malmquist index (output orientation)	Malmquist index=Efficiency Change× Frontier-Shift Where, Efficiency change=$\frac{D_0^t(x^t,y^t)}{D_0^s(x^s,y^s)}$; Frontier-shift=$[\frac{D_0^s(x^t,y^t)}{D_0^t(x^t,y^t)}\frac{D_0^s(x^s,y^s)}{D_0^t(x^s,y^s)}]^{1/2}$ In the frontier-shift formula, $D_0^s(x^t, y^t)$ measures the efficiency at time *t* against the frontier at time *s*; similarly, $D_0^t(x^s, y^s)$ measures the efficiency at time *s* against the frontier at time *t*. Both $D_0^s(x^t, y^t)$ and $D_0^t(x^s, y^s)$ are also called cross-efficiencies.		
Interpretation	Efficiency change: >1: positive; =1: no change; <1: negative	Frontier-shift /technical change: >1: forward/positive; =1: no change; <1: backward/negative	Malmquist index (m0): >1: increase in productivity; =1: no change; <1: decrease in productivity

As mentioned in section 2.4.3, a VRS DEA model may cause bias in the calculation of productivity change (Grifell-Tatjé and Lovell, 1995), the Malmquist indices in this study are calculated using CRS models.

To examine the changes in school efficiencies since the introduction of NCEA, only the schools with data in both 2004 and 2011 were included for the calculation of Malmquist index. As such, 277 schools were retained in the full sample whilst 199 schools were retained in the pure secondary school sample.

As a starting point to analyse the Malmquist indices for school efficiencies, it is essential to examine if there were any changes in model variables over the period (see Table 4.22). Of note, for the purpose of comparison or modelling the Malmquist index, other operational expenditures in 2011 should exclude the amount for the 'use of land and buildings' (UoLB), given that this item was introduced in 2007, as a result of a change in accounting policy to recognise the non-monetary MoE grant relating to land and buildings. To obtain a fair and reasonable comparison result, measures for school efficiencies in 2004 and 2011 must be on the same basis.

Table 4.22 Change (%) in variables from 2004 to 2011

Change (%)[70] in variables from 2004 to 2011	Full sample (n=277)		Pure secondary school sample (n=199)	
Base year: 2004	Mean	St. Dev.	Mean	St. Dev.
Indexed teachers' salaries	16%	13%	14%	16%
Indexed other expenditures[71]	10%	20%	14%	15%
Roll	1%	5%	1%	8%
# NCEA Level 1 passes	7%	2%	7%	1%
# NCEA Level 2 passes	18%	1%	18%	-1%
# NCEA Level 3 passes	48%	22%	44%	20%
# UE	39%	18%	36%	22%
# NCEA L3 M& E	69%	53%	60%	47%
SES	-2%	-2%	-2%	-2%

It appears that for both samples, the change rates of teachers' salaries and other operating expenditures are greater than that of roll. These findings can explain the decline in school cost efficiency from 2004 to 2011. In contrast, the average change rates of NCEA Level 1-3 passes are greater than that of roll, indicating a growth in school academic efficiency. Likewise, the average rates of change in the number of UE and number of NCEA L3 M&E endorsements are higher than those of NCEA Level 1-3 passes, showing an increase in school academic effectiveness. Similarly, the average rates of change in UE passes and number of NCEA L3 M&E are also greater than those of teachers' salaries and other operating expenditures, suggesting an improvement in school overall efficiency since the introduction of the NCEA system. This is due to the effect of a policy change for the improvements to NCEA during 2007-2008, which aimed to support teachers' assessment practice, improving student motivation via introducing the M&E endorsement of NCEA certificates and publishing more NCEA statistics (New Zealand Qualifications Authority, 2015). As a consequence, school management and teachers developed and refined teaching strategies, curriculum framework and assessment system so as to maximize school academic outcomes from their limited resources.

Based on the formulas specified by Färe et.al (1989, 1994), the Malmquist indices for all types of efficiencies were calculated using the categorical CRS models under output orientation. The categorical models were selected to incorporate the effect of SES decile on school efficiencies. In this regard, the indices presented for each decile reflect the

[70] Change (%) is computed as the difference between 2004 and 2011 divided by 2004 value.
[71] This amount excludes the value of 'use of lands and buildings'.

productivity change for schools in that decile relative to all schools sitting in the same or lower decile.

However, as mentioned earlier, there is a significant change in school decile during the test period. For the full sample, 127 out of 277 schools in 2004 have different decile from that in 2011. To address this issue for the calculation of Malmquist indices, schools in 2004 and 2011 were regarded as individual unit while being categorized into their decile group in the starting and ending years. For example, Bay of Islands College was a decile two school in 2004, then its decile ranking changed to one in 2011. As such, to calculate its efficiency change[72] from 2004 to 2011 using categorical CRS model, its efficiency score in 2004 ($D_0^s(x^s, y^s)$) was produced in the sample comprised of decile one and decile two schools, whereas that in 2011, $D_0^t(x^t, y^t)$ was calculated in the category of decile one schools only.

Further, to compute the frontier shift for a specific school, two additional distance functions being $D_0^t(x^s, y^s)$ and $D_0^s(x^t, y^t)$ are needed to address the differences between time *s* and *t* under the categorical CRS model. Taking Bay of Islands College as an example, $D_0^t(x^s, y^s)$ measures the efficiency for 2004 against 2011 decile one frontier, whereas $D_0^s(x^t, y^t)$ measures the efficiency for 2011 against 2004 decile two frontier.

For each school in the sample, the beginning and ending school years (i.e., 2004 and 2011) were included in the categorical DEA model to compute the cross-efficiencies, and each school year was categorized by the decile in the corresponding year. Taking Bay of Islands College as an example, its 2004 school year was classified in decile two group whereas 2011 school year was categorized in decile one group. These two school years were included in the sample for the calculation of each cross-efficiency using the categorical CRS model. Under the output orientation mode for each efficiency (i.e., cost efficiency, academic efficiency, academic effectiveness and overall efficiency), both $D_0^t(x^s, y^s)$ and $D_0^s(x^t, y^t)$ were produced using the pooled data comprised by both 2004 and 2011 school years for all the schools selected for the Malmquist index analysis. With respect to the calculation for $D_0^t(x^s, y^s)$, for 2004 school years, both inputs and outputs were exactly the same as those in the original model, whereas for 2011 school years, inputs were the original but outputs were a proportion of the

[72] In this study, efficiency change for each type of efficiency (i.e., cost efficiency, academic efficiency, academic effectiveness and overall efficiency) is calculated.

original. Following a rule of thumb, 0.1 was taken as the proportion ratio on the original outputs to run the DEA model. As such, $D_0^t(x^s, y^s)$ scores were obtained through multiplying ten by the DEA results for 2011 school years; similarly, to compute $D_0^s(x^t, y^t)$ scores, for 2011 school years, both inputs and outputs were exactly the same as those in the original model, whereas for 2004 school years, inputs were the original but outputs were 0.1 of the original. After running the DEA model, $D_0^s(x^t, y^t)$ scores were acquired through multiplying ten by the DEA results for 2004 school years.

However, with respect to the calculation of cross-efficiencies, the above approach does not incorporate the possible frontier shift caused by the change in decile during the test period, which might distort the cross-efficiency scores as well as the Malmquist indices. To examine the impact of change in decile on the results, an additional calculation for Malmquist indices and their decomposition scores was undertaken for the full sample[73] excluding schools with decile changed between 2004 and 2011. As a result, there were 150 schools retained in the full sample after the exclusion. The *t*-test results suggest that for any one of cost efficiency, academic efficiency, academic effectiveness and overall efficiency, there are no significant differences in the mean Malmquist index and its mean decomposition results for each decile between the initial sample and the test sample, which only contains schools with constant decile. Therefore, in this study, the effect of decile change is ignored in the process of calculating Malmquist indices and its decomposition results.

Once all the distance functions were obtained, the frontier shift for an individual school was produced via the equation listed in Table 4.21 (see also Equation 2.6).

Tables 4.23-4.26 show the mean Malmquist indices for both samples, together with their decomposition into efficiency change and frontier-shift in terms of cost efficiency, academic efficiency, academic effectiveness and overall efficiency. Tables A12.1-A12.4 in Appendix XII detail the percentages of schools with increase or decrease in efficiency change, frontier-shift and overall productivity change. All these results are tabulated by decile in 2004.[74]

[73] Since the pure secondary school sample is approximately a subset of the full sample, this test is only conducted on the full sample.

[74] Decile in 2004 was selected because this analysis aims to look at the changes in school performance efficiencies since NCEA was introduced in 2004.

According to Table 4.23, for the full sample, the mean Malmquist index for cost efficiency in each decile is less than one, due to the decline in both relative efficiency and technical innovation (frontier-shift). The decline in cost efficiency is consistent with the fact that school expenditures increased much more than school rolls over the period. The backward shift of frontier is also in line with the growth of teachers' salaries during the test period as reported by the MoE: the policy for teachers' pay had been adjusted over time and resulted in the increase of salaries[75] at an average of more than 10% per year from 2004 to 2011 (Ministry of Education, 2013a). The standard deviations for frontier-shift are overall smaller than those for efficiency change. The less variability of frontier-shift suggests a similar proportion of increases in school expenditures for the majority of schools.

Table 4.23 Malmquist index decomposition for cost efficiency

Full sample: cost efficiency							
		Efficiency change		**Frontier shift**		**Malmquist index**	
SES[76]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	10	0.96	0.10	0.94	0.04	0.90	0.10
2	25	0.95	0.11	0.85	0.12	0.81	0.16
3	25	1.01	0.12	0.87	0.05	0.87	0.09
4	31	0.99	0.09	0.90	0.03	0.88	0.08
5	33	0.94	0.12	0.92	0.04	0.86	0.10
6	41	0.95	0.11	0.96	0.04	0.91	0.11
7	29	0.92	0.09	0.97	0.04	0.89	0.09
8	30	0.95	0.11	1.0	0.11	0.96	0.18
9	27	0.94	0.08	0.96	0.03	0.90	0.08
10	26	0.94	0.09	0.95	0.03	0.89	0.07
Total	277	0.95	0.11	0.93	0.07	0.89	0.12
Pure secondary school sample: cost efficiency							
		Efficiency change		**Frontier shift**		**Malmquist index**	
SES[77]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	6	0.97	0.06	1.06	0.03	1.04	0.06
2	21	0.95	0.11	0.90	0.14	0.86	0.17
3	20	1.03	0.11	0.91	0.02	0.93	0.09
4	22	1.0	0.10	0.91	0.03	0.90	0.08
5	24	0.95	0.12	0.92	0.02	0.86	0.11
6	29	0.94	0.08	0.93	0.02	0.88	0.07
7	24	0.94	0.09	0.93	0.02	0.88	0.08
8	18	0.94	0.08	0.92	0.02	0.87	0.07
9	21	0.97	0.08	0.93	0.02	0.90	0.07
10	14	0.96	0.04	0.93	0.02	0.90	0.03
Total	199	0.97	0.09	0.92	0.05	0.89	0.10

[75] This result has not been adjusted for the inflation effect.
[76] This column records school decile in 2004.
[77] Ibid.

With regard to the productivity change in cost efficiency for the pure secondary school sample, similar findings were obtained except for the decile one schools, which had improvements in both efficiency and technical innovation over the period. However, interpretation for this phenomenon must be made with caution due to the sample bias, given that only six schools were included in decile one for the pure secondary school sample.

Table A12.1 in Appendix XII show that for both samples, more than 60% of schools suffered a decrease in cost efficiency (FS: 69%; PS: 64.8%) and an even larger proportion of schools had a backward frontier-shift (FS: 92.4%; PS: 97%). As a result, there were about 88% of schools with an overall decline in cost efficiency over the period. Accordingly, there was overall a negative change in productivity for cost efficiency since the introduction of NCEA, in particular, the majority of schools had a backward frontier-shift over the period.

In contrast with the cost efficiency, Table 4.24 displays that for both samples, the mean Malmquist indices of academic efficiency for decile 1-9 are all greater than one, mainly due to the positive efficiency change in most decile 1-9 schools over the period. This is a consequence of the significant increase in NCEA passes relative to the stable rolls over time. The frontier moved forward for decile 1-9 schools in the full sample and for decile 1-7 schools in the pure secondary school sample. This suggests that most schools, especially those in lower and medium decile, achieved improvements in not only relative efficiency but also technical innovation since the introduction of NCEA. However, results show that decile ten schools in both samples had unchanged efficiency but experienced a backward frontier-shift over the period, this is partially because some decile ten schools focus on the CIE system (e.g., Auckland Grammar School and Macleans College), they therefore achieve less NCEA passes. Results from Table A12.2 in Appendix XII show that, only 34.6% decile ten schools in the full sample and none in the pure secondary school sample had forward frontier-shift for academic efficiency over the period. Indeed, the catch-up of lower deciles under the NCEA system demonstrates that NCEA levels the playing field for schools with different SES decile, through benefiting decile 1-9 rather than decile ten schools in terms of achieving more national qualifications.

Table 4.24 Malmquist index decomposition for academic efficiency

Full sample: academic efficiency							
		Efficiency change		Frontier shift		Malmquist index	
SES[78]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	10	1.16	0.47	1.69	0.36	1.87	0.54
2	25	1.15	0.26	1.22	0.22	1.43	0.56
3	25	1.07	0.21	1.20	0.09	1.29	0.27
4	31	1.12	0.21	1.13	0.08	1.25	0.22
5	33	1.08	0.18	1.10	0.08	1.19	0.21
6	41	1.04	0.16	1.12	0.08	1.16	0.16
7	29	1.03	0.19	1.10	0.04	1.14	0.20
8	30	1.03	0.16	1.06	0.05	1.09	0.16
9	27	1.04	0.12	1.04	0.03	1.08	0.11
10	26	1.00	0.17	0.99	0.05	0.99	0.19
Total	277	1.06	0.21	1.13	0.17	1.20	0.31
Pure secondary school sample: academic efficiency							
		Efficiency change		Frontier shift		Malmquist index	
SES[79]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	6	1.07	0.24	1.06	0.07	1.13	0.22
2	21	1.13	0.26	1.11	0.07	1.25	0.27
3	20	1.15	0.26	1.12	0.06	1.29	0.29
4	22	1.11	0.19	1.06	0.07	1.18	0.22
5	24	1.10	0.17	1.03	0.08	1.13	0.17
6	29	1.08	0.16	1.03	0.06	1.11	0.17
7	24	1.09	0.16	1.03	0.04	1.12	0.17
8	18	1.11	0.15	0.98	0.05	1.09	0.18
9	21	1.06	0.11	0.97	0.03	1.03	0.12
10	14	1.0	0.23	0.90	0.04	0.90	0.23
Total	199	1.09	0.19	1.03	0.08	1.13	0.22

Results from Table A12.2 also show that for both samples, more than 60% of schools enjoyed an increase in academic efficiency (FS: 60.6%; PS: 62.8%); there were 91.7% schools in the full sample and 68.3% schools in the pure secondary school sample having forward frontier-shift. As a consequence, there were 79.1% schools in the full sample and 69.3% in the pure secondary school sample, having an improvement in productivity for academic efficiency over the period. Given that the only difference between the two samples is the inclusion of expanded schools in the full sample, this finding indicates some exogenous factors might relate to the productivity change in academic efficiency as many expanded schools are integrated schools with relatively small rolls. The relationships between productivity change and exogenous factors will be examined in section 4.6.2.

[78] Ibid.
[79] Ibid.

Table 4.25 shows that for both samples, the mean Malmquist index of academic effectiveness for all deciles is greater than one, predominantly due to the improvement in relative efficiency in academic effectiveness. The mean frontier-shift of the Malmquist index decomposition indicates no change in technical innovation to obtain academic effectiveness over the period. However, the split results by decile show that, except for decile one,[80] other lower decile schools (i.e. decile 2-4) are less capable of achieving a forward frontier-shift than schools with higher decile (i.e., decile 8-10). In fact, decile 2-4 schools had backward frontier-shifts whilst decile 8-10 schools achieved forward frontier movements, with decile 5-7 schools having frontiers almost unmoved over the period. This is consistent with the expectation for high decile schools, in particular, decile ten schools, having more capability to transfer academic outputs (i.e., NCEA passes) into academic outcomes (i.e., UE passes and NCEA L3 M&E endorsements).

[80] To be discussed later in this section.

Table 4.25 Malmquist index decomposition for academic effectiveness

Full sample: academic effectiveness							
		Efficiency change		Frontier shift		Malmquist index	
SES[81]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	10	1.25	0.41	1.04	0.18	1.29	0.41
2	25	1.15	0.37	0.85	0.16	0.99	0.43
3	25	1.04	0.22	0.92	0.09	0.96	0.24
4	31	1.08	0.26	0.98	0.11	1.06	0.27
5	33	1.07	0.33	1.01	0.09	1.07	0.29
6	41	1.16	0.38	1.01	0.09	1.18	0.44
7	29	1.01	0.12	1.00	0.11	1.01	0.16
8	30	1.06	0.15	1.02	0.09	1.08	0.17
9	27	1.03	0.14	1.09	0.15	1.12	0.20
10	26	1.06	0.11	1.12	0.12	1.19	0.18
Total	277	1.08	0.27	1.00	0.13	1.09	0.30
Pure secondary school sample: academic effectiveness							
		Efficiency change		Frontier shift		Malmquist index	
SES[82]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	6	1.09	0.44	1.27	0.23	1.35	0.49
2	21	1.16	0.39	0.84	0.20	0.99	0.50
3	20	1.08	0.21	0.90	0.08	0.97	0.23
4	22	1.04	0.24	0.92	0.12	0.96	0.23
5	24	1.01	0.20	1.00	0.10	1.00	0.19
6	29	1.15	0.41	1.00	0.11	1.16	0.49
7	24	0.99	0.12	1.02	0.09	1.01	0.15
8	18	1.04	0.10	1.06	0.10	1.11	0.16
9	21	0.99	0.13	1.11	0.13	1.10	0.19
10	14	1.02	0.10	1.11	0.13	1.13	0.17
Total	199	1.06	0.26	1.00	0.15	1.06	0.31

Statistics in Table A12.3 in Appendix XII show that, more than 50% of schools in both samples enjoyed an improvement in efficiency to achieve academic effectiveness (FS: 55.6%; PS: 50.3%), while less than 50% of schools in both samples had a forward frontier-shift (FS: 43%; PS: 49.2%). However, there were still over 50% of schools that attained positive change in productivity for academic effectiveness. Consistent with the above discussion, higher decile especially decile ten schools, show a better ability to achieve the improvement in technical innovation, given more than 53% of decile 8-10 schools had a forward frontier-shift for academic effectiveness. Therefore, there is overall a positive productivity change in academic effectiveness since the introduction of NCEA and higher decile schools show better capability of achieving the improvement in technical innovation for academic effectiveness.

[81] Ibid.
[82] Ibid.

Likewise, as shown in Table 4.26, for both samples, the mean Malmquist index of overall efficiency for all deciles is greater than one, mainly due to the forward frontier-shifts, given the mean frontier-shift of Malmquist index decomposition in each decile is higher than one. Apart from decile two, schools in all the other deciles achieved improvement in relative efficiency for overall efficiency, showing the overall increase in academic outcomes generated from the budgeted expenditures. However, the relatively high standard deviations suggest a wide variation of efficiency change in each decile, in particular, the lowest deciles (i.e., decile 1-3). This is in consistent with the wide variety in the number of NCEA L3 M&E endorsements, where the variety is more identifiable in the lowest deciles.

Table 4.26 Malmquist index decomposition for overall efficiency

Full sample: overall efficiency							
		Efficiency change		**Frontier shift**		**Malmquist index**	
SES[83]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	10	1.98	1.71	1.91	0.39	3.60	2.84
2	25	0.95	0.74	1.58	0.30	1.52	1.26
3	25	1.18	0.64	1.28	0.21	1.50	0.78
4	31	1.37	0.64	1.22	0.11	1.69	0.84
5	33	1.35	1.29	1.22	0.17	1.68	1.79
6	41	1.41	0.95	1.24	0.12	1.74	1.16
7	29	1.03	0.39	1.38	0.24	1.41	0.52
8	30	1.08	0.33	1.49	0.34	1.61	0.60
9	27	1.06	0.34	1.29	0.15	1.37	0.44
10	26	1.08	0.33	1.14	0.10	1.23	0.40
Total	277	1.22	0.80	1.33	0.27	1.62	1.17
Pure secondary school sample: overall efficiency							
		Efficiency change		**Frontier shift**		**Malmquist index**	
SES[84]	N	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
1	6	1.22	0.89	2.65	0.35	3.04	1.94
2	21	0.93	0.75	1.57	0.31	1.50	1.29
3	20	1.18	0.55	1.16	0.10	1.37	0.67
4	22	1.19	0.47	1.15	0.04	1.37	0.56
5	24	1.23	0.49	1.10	0.04	1.35	0.54
6	29	1.29	0.58	1.10	0.04	1.41	0.64
7	24	1.19	0.36	1.10	0.03	1.31	0.43
8	18	1.17	0.41	1.16	0.06	1.37	0.54
9	21	1.09	0.35	1.26	0.10	1.38	0.44
10	14	1.07	0.38	1.10	0.10	1.18	0.43
Total	199	1.16	0.52	1.23	0.32	1.42	0.77

[83] Ibid.
[84] Ibid.

Results of Table A12.4 in Appendix XII demonstrate that more than 50% of schools in both samples enjoyed an increase in the relative efficiency for overall efficiency (FS: 50.9%; PS: 55.8%). Meanwhile, there were more than 98% of schools had forward frontier-shift. As a consequence, more than 75% of schools achieved an improvement in productivity for overall efficiency over time (FS: 81.9%; PS: 75.4%). Accordingly, there is overall a positive productivity change in overall efficiency since the introduction of NCEA, primarily due to the improvements in technical innovation to achieve academic outcomes from limited resources.

It is worth noting that with regards to all the efficiencies, decile one schools achieved more improvements in both relative efficiency and technical innovation compared with other lower decile schools (i.e., decile 2-3). While decile one schools are likely to receive higher efficiency scores than other lower decile schools due to the smallest sample size where the categorical model is applied, the targeted assistance provided by the MoE and ERO to decile one schools does appear to have resulted in significant improvement in academic performance over the period. A further review of the data of decile one schools revealed that most of them obtained a positive change in productivity due to their significant progress in academic outputs, while only the minority achieved the improvement as a result of the reduction in school roll. Nevertheless, schools in other low deciles did not accomplish parallel improvements.

4.5.5 Summary on DEA application

The comparison of efficiency scores between the standard model and categorical model suggests that results from the categorical model better reflect New Zealand school performance than the standard model, because the categorical model adjusts the effect of student social-economic background which has significant impact on school academic outputs and outcomes.

For both samples, the mean scale efficiency for all types of efficiency under both categorical and standard models are over 88%, suggesting that the majority of New Zealand secondary schools are running nearly at the most productive sizes.

Results from both standard VRS and categorical VRS models demonstrate that all types of efficiency scores are positively related to SES decile. Even though scores produced by the categorical model are flatter among different decile levels compared to those from the standard model, they are overall positively correlated to decile. Meanwhile, efficiency scores are also positively related to school rolls as a result of large schools normally having higher decile. These findings indicate that large and high decile schools generally outperform small and low decile schools in terms of all the efficiencies.

Comparison of mean efficiency scores by school characteristics indicates that some school characteristics are associated with different types of efficiency. Typically, in terms of academic efficiency and effectiveness, girl schools outperform boy schools and their co-educational peers whilst integrated schools perform better than non-integrated schools. For the full sample, expanded schools outperform pure secondary schools in terms of all types of efficiency due to their higher mean decile. These findings are also consistent with the numbers of fully efficient schools categorized by school characteristics with respect to the corresponding efficiencies.

To examine whether the network DEA model can better reflect school performance compared to a single DEA model, comparison was made to look at the differences between overall efficiency and sub-efficiencies. It appears that compared to the network DEA model, the overall efficiency model evaluates school performance by a single composite score, which hides some critical information for performance evaluation and improvement, both for schools classified as efficient and inefficient.

In general, results for 2004-2006 are similar to those for 2009-2011 with some quantitative differences in the strength of correlation between the efficiency scores and the model/characteristic variables. However, comparison of efficiency scores within each period reveals that academic efficiency and effectiveness in the latest year of each period increased while cost efficiency for 2004-2006 decreased over the years due to the increased Government grants spending on school NCEA-related activities.

The examination of returns to scale (RTS) suggests that under the categorical VRS model, no schools were identified suffering DRS. This indicates that New Zealand secondary school performance could be improved by enlarging school size.

The results of Malmquist index and its decomposition show that since the introduction of NCEA, there are a number of changes in relative efficiency, frontier-shift and overall productivity with respects to cost efficiency, academic efficiency, academic effectiveness and overall efficiency. First, there is overall a decline in productivity for cost efficiency as a result of the backward frontier-shift, because teachers' salaries increased significantly while school rolls remained stable during the test period. Second, there is overall an improvement in productivity for academic efficiency since the introduction of NCEA and most schools enjoyed both an increase in relative efficiency and a forward frontier-shift. However, the productivity of decile ten schools remained unchanged, this is likely because some decile ten schools focused more on the CIE system thus received less NCEA passes compared to their peers, Third, there is overall an increase in productivity for academic effectiveness during the test period mainly due to the growth in relative efficiency. Detailed examination reveals that higher decile schools have better capability of achieving the improvement in technical innovation for academic effectiveness. Finally, there is a positive productivity change in overall efficiency on average primarily due to the improvement in technical innovation to achieve academic outcomes from limited resources. Typically, compared with decile two and decile three schools, decile one schools appeared to have much more improvements in both relative efficiency and technical innovation for all types of efficiency. This suggests that the targeted assistance from the MoE in improving their academic outcomes was successful.

4.6 Regression analysis

As discussed earlier, some exogenous factors (i.e., school ownership etc.) excluded from the network DEA model might be associated with school performance at different levels. For example, integrated schools are expected to have better academic outputs than non-integrated schools; single-sex schools outperform co-educational schools in terms of academic achievements. To examine the relationships between efficiency scores/productivity changes and each exogenous factor of interest, several ordinary least square (OLS) regressions [85] were developed to check if the co-efficient on the corresponding factor/variable is significant while controlling the relationships with other variables. A significant co-efficient on a variable explains its relationship with the corresponding efficiency or productivity change.

4.6.1 Regressions on level of efficiency

To examine the relationships between each type of efficiency and exogenous variables such as ownership, gender and Mäori density, an OLS regression was developed as follows:

Equation 4.1 Regression equation for efficiency scores

$$\boldsymbol{E_i^t = \alpha + \beta_1\, O^t + \beta_2\, C^t + \beta_3\, R^t + \beta_4\, M^t + \varepsilon}$$

Where *E* is the efficiency scores under the categorical VRS model; *i* represents anyone of cost efficiency, academic efficiency, academic effectiveness or overall efficiency; *t* represents the year; *O* is the ownership indicator, where the value 1 means integrated and 0 otherwise; *C* is the co-educational school indicator, where the value 1 refers to co-educational schools and 0 otherwise; *R* is standardized school roll [86] and *M* is the indicator of Mäori density，where the value 1 represents schools with at least 30% Mäori students [87] and 0 means schools with less than 30% Mäori roll.

[85] Given the bounded nature of DEA efficiency scores, Tobit regression is a widely-used method to handle the truncation issue. However, research reveals that results from OLS regression can also explain DEA efficiency scores well just as those from Tobit regression (Hoff, 2006). Therefore, OLS regression approach is adopted in this study.

[86] Since other independent variables in the regression are between 0 and 1, roll is standardized into a range between 0 and 1 to ensure a distinguishable coefficient attached to it. Specifically, standardized roll= (roll-minimum roll)/roll range.

[87] 30% is the 75% percentile of the rate of Maori students for both samples during 2009-2011. The upper quartile for both samples in 2004-2006 is 27%, therefore 30% was selected to control Maori density in the regression model.

Since statistics suggest that single-sex [88] schools generally have distinguishable performance from co-educational schools in terms of academic efficiency, academic effectiveness and overall efficiency (Appendix XIII), the indicator for co-educational school is included in the regression formula. Moreover, statistics also show that Mäori students' pass rate is overall lower than that of other ethnic groups (see Appendix XIV). Therefore, M is included in the model to test the relationship between Mäori density and efficiency level. However, because Mäori rate is associated with school roll that is found significantly related to most types of efficiency, standardized roll is included in the regression to control for this effect.

Based on the results detailed in Tables 4.12 and 4.14, positive regression coefficient was expected on the ownership variable in terms of academic efficiency, academic effectiveness and overall efficiency. Moreover, a negative coefficient was expected on co-educational school variable in terms of academic efficiency, academic effectiveness and overall efficiency. A negative coefficient was also expected on Mäori30% variable in terms of academic efficiency. The sign of the coefficient on Mäori30% for academic effectiveness cannot be predicted properly because academic effectiveness measures the efficiency of converting NCEA passes to UE passes and the number of NCEA L3 M&E endorsements. A high level of academic effectiveness can be obtained in the case that the number of NCEA passes is very small whilst the number of UE passes or the number of NCEA L3 M&E endorsements is not zero.

The regression result regarding cost efficiency for the full sample in 2009-2011 shows that cost efficiency is positively related to Mäori30% variable. This is unexpected because almost half of Mäori30% schools have low decile (i.e., decile 1-3), and the Government provides increased funding into these schools while most of them have roll size relatively constant in each test period. A further analysis revealed that for 2009-2011, Mäori30% schools in decile 1-3 have higher cost per student as a result of lower mean roll compared to non-Mäori30% schools. In contrast, Mäori30% schools in decile 4-6 have lower cost per student due to a higher mean roll than non-Mäori30% schools. Therefore, it is expected that for different decile groups, the coefficients on Mäori30%

[88] The initial regression model included girl school indicator as an independent variable, together with co-educational school indicator. However, it is evident that girl school indicator is significantly correlated to co-educational school indicator and this causes distortion against the coefficient on co-educational school variable. Since the majority of New Zealand secondary schools are co-educated, girl school indicator is finally excluded from the model to ensure the robustness of the results related to co-educational schools.

indicator will have different signs and significances for cost efficiency and other efficiencies.

An additional review of the dataset showed that for 2009-2011, Mäori30% schools in both samples are only in decile 1-6; similarly, for 2004-2006, the majority of Mäori30% schools in both samples are decile 1-6 schools with only two schools in decile seven. As such, regression analysis was undertaken for three decile groups, being decile 1-3, decile 4-6 and decile 7-10. The regression results for the two samples in both test periods are detailed in Tables 4.27- 4.30.

For the full sample during 2009-2011, there are a number of findings outlined as follows:

- For integrated schools, both cost efficiency and academic efficiency are positively related to those in decile 1-6; academic effectiveness is positively correlated to integrated schools in decile 7-10; overall efficiency is positively related to integrated schools at all decile levels;
- For co-educational schools, both academic efficiency and overall efficiency are negatively associated with co-educational schools at all decile levels; cost efficiency is negatively related to co-educational schools in decile 1-3 and academic effectiveness is negatively correlated to co-educational schools in decile 1-3 and decile 7-10 groups;
- For Mäori30% schools, cost efficiency is negatively associated with those in decile 1-3 group but positively related to those in decile 4-6 group. During the test period of 2009-2011, Mäori30% schools in decile 1-3 group have smaller mean size and higher mean cost per student, compared to non-Mäori30% schools in the same decile group. As a consequence, Mäori30% schools are less efficient in terms of cost efficiency compared to non-Mäori30% schools. Therefore, a positive change in cost efficiency for schools in decile 1-3 is more associated with non-Mäori30% schools and this explains the negative relationship between cost efficiency and Mäori30% schools in decile 1-3 group. In contrast, Mäori30% schools in decile 4-6 group have larger mean roll and lower mean cost per student compared to non-Mäori30% schools, therefore, cost efficiency for decile 4-6 group is positively associated with Mäori30% schools. With regards to the other efficiencies, academic effectiveness is negatively associated with Mäori30% schools in decile 1-3. There is no evidence showing that either academic efficiency or overall efficiency is associated with Mäori30% schools;

- For school roll, both cost efficiency and overall efficiency are positively associated with school roll at all decile levels; academic efficiency is positively related to roll for decile 1-6 schools; academic effectiveness is positively correlated to roll for decile 4-10 schools.

Regression results for the pure secondary school sample in 2009-2011 are similar to those for the full sample, except for the following differences:

- There is no evidence showing that any efficiencies are associated with integrated schools in the decile 7-10 group; academic effectiveness is positively related to integrated schools in decile 4-6 but not 7-10. However, the smaller number of integrated schools and the weak significance suggests that the findings from the full sample are more convincing, given that a large proportion of integrated schools are expanded secondary schools (i.e., years 7-15 schools);
- There is no evidence suggesting that cost efficiency is related to Mäori30% schools in decile 1-3; unlike the result for the full sample, overall efficiency is significantly and negatively associated with Mäori30% schools in decile 4-6 group. The differences from the full sample are due to the exclusion of expanded schools (years 7-15) in the pure secondary school sample, which results in the change in the distribution of Mäori30% schools in each decile group.

Compared to the 2009-2011 period, results for the full sample in 2004-2006 are similar except for the following features:

- For integrated schools, academic efficiency is positively associated with integrated schools at all decile levels, rather than just related to decile 1-6 group; there is no evidence showing that academic effectiveness is associated with integrated schools;
- For co-educational schools, both academic efficiency and overall efficiency are negatively related to co-educational schools in decile 4-10 rather than at all decile levels; academic effectiveness is negatively related to co-educational schools in decile 4-10, rather than decile 1-3 and decile 7-10; cost efficiency is negatively related to co-educational schools in decile 4-6, rather than decile 1-3;
- There is no evidence showing that any efficiencies are related to Mäori30% schools in decile 1-3; however, academic efficiency is negatively correlated to Mäori30% schools in decile 4-6;

- For school roll, in addition to cost efficiency and overall efficiency, academic efficiency and academic effectiveness are also positively associated with the standardized roll for schools at all decile levels.

For the pure secondary school sample in 2004-2006, a number of different features from the full sample are outlined as below:

- Integrated schools: cost efficiency is positively related to integrated schools at all decile levels rather than just for decile 1-6 integrated schools;
- Co-educational schools: both cost efficiency and academic effectiveness are negatively associated with co-educational schools at all decile levels rather than some decile groups.

Table 4.27 Full sample: regression results for all types of efficiency in 2009-2011

		Cost efficiency			Academic efficiency			Academic effectiveness			Overall efficiency		
		Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.
Decile 1-3 group N=198	Intercept	0.765		0***	0.781		0***	0.903		0***	0.596		0***
	Ownership	0.086	0.239	0***	0.147	0.39	0***	0.064	0.127	0.111	0.302	0.368	0***
	Co-educated	-0.048	-0.133	0.021**	-0.065	-0.174	0.006**	-0.08	-0.16	0.025**	-0.17	-0.207	0.001**
	Mäori30%	-0.037	-0.144	0.034**	-9.11E-05	0	0.996	-0.07	-0.194	0.021**	-0.068	-0.115	0.118
	Std. Roll	0.497	0.524	0***	0.368	0.371	0***	0.075	0.057	0.493	0.645	0.298	0***
	F-value	34.763***			19.929***			6.061***			23.009***		
	Adjusted R^2	0.407			0.278			0.093			0.309		
Decile 4-6 group N=331	Intercept	0.673		0***	0.821		0***	0.678		0***	0.53		0***
	Ownership	0.057	0.148	0.002**	0.067	0.184	0.002**	0.045	0.089	0.138	0.18	0.273	0***
	Co-educated	-0.016	-0.06	0.192	-0.03	-0.114	0.043**	-0.006	-0.017	0.765	-0.077	-0.164	0.002**
	Mäori30%	0.05	0.189	0***	-0.008	-0.032	0.554	-0.008	-0.022	0.695	-0.035	-0.077	0.13
	Std. Roll	0.414	0.616	0***	0.078	0.121	0.031**	0.132	0.147	0.011**	0.355	0.308	0***
	F-value	49.007***			6.011***			2.232*			21.467***		
	Adjusted R^2	0.368			0.057			0.015			0.199		
Decile 7-10 group N=336	Intercept	0.651		0***	0.891		0***	0.692		0***	0.552		0***
	Ownership	0.012	0.051	0.342	0.009	0.038	0.597	0.063	0.233	0.001**	0.082	0.201	0.003**
	Co-educated	0.007	0.031	0.489	-0.024	-0.106	0.084*	-0.046	-0.179	0.002**	-0.079	-0.203	0***
	Mäori30%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Std. Roll	0.394	0.7	0***	-0.019	-0.034	0.609	0.242	0.379	0***	0.332	0.345	0***
	F-value	93.784***			2.251*			18.9***			17.495***		
	Adjusted R^2	0.454			0.011			0.138			0.129		

Dependent variable: VRS efficiency score; significance: ****p*<0.001, ***p*<0.05; * *p*<0.1 (1-tailed).

Note: with regards to a regression equation with adjusted R^2less than 0.1, the low explanatory power suggests there are other unidentified exogenous factors related to the efficiency.

Table 4.28 Pure secondary school sample: regression results for all types of efficiency in 2009-2011

		Cost efficiency			Academic efficiency			Academic effectiveness			Overall efficiency		
		Beta	**Std. Beta**	**Sig.**	**Beta**	**Std. Beta**	**Sig.**	**Beta**	**Std. Beta**	**Sig.**	**Beta**	**Std. Beta**	**Sig.**
Decile 1-3 group N=153	Intercept	0.845		0***	0.827		0***	0.927		0***	0.607		0***
	Ownership	0.108	0.305	0***	0.143	0.315	0***	0.037	0.063	0.48	0.336	0.327	0***
	Co-educated	-0.036	-0.138	0.055*	-0.068	-0.204	0.007**	-0.058	-0.132	0.118	-0.197	-0.26	0***
	Mäori30%	-0.016	-0.079	0.32	-0.011	-0.044	0.597	-0.057	-0.171	0.068*	-0.06	-0.104	0.185
	Std. Roll	0.31	0.462	0***	0.291	0.337	0***	-0.016	-0.014	0.883	0.709	0.362	0***
	F-value	16.875***			12.036***			2.021*			18.185***		
	Adjusted R^2	0.295			0.225			0.026			0.311		
Decile 4-6 group N=240	Intercept	0.757		0***	0.837		0***	0.67		0***	0.516		0***
	Ownership	0.067	0.157	0.008**	0.112	0.206	0.002**	0.109	0.154	0.024**	0.268	0.291	0***
	Co-educated	-0.013	-0.064	0.254	-0.047	-0.179	0.005**	0.003	0.01	0.88	-0.106	-0.239	0***
	Mäori30%	0.025	0.12	0.031**	-0.008	-0.031	0.619	-0.029	-0.085	0.185	-0.046	-0.104	0.056*
	Std. Roll	0.284	0.558	0***	0.093	0.145	0.025**	0.169	0.2	0.002**	0.483	0.442	0***
	F-value	25.253***			6.618***			3.844**			28.966***		
	Adjusted R^2	0.289			0.086			0.045			0.319		
Decile 7-10 group N=222	Intercept	0.775		0***	0.919		0***	0.708		0***	0.672		0***
	Ownership	0.011	0.05	0.48	0.025	0.079	0.327	0.016	0.046	0.549	0.013	0.027	0.735
	Co-educated	-0.019	-0.112	0.079*	-0.012	-0.052	0.46	-0.046	-0.178	0.009**	-0.085	-0.225	0.001**
	Mäori30%	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	Std. Roll	0.216	0.52	0***	-0.082	-0.142	0.063*	0.232	0.361	0***	0.169	0.182	0.016**
	F-value	23.37***			3.468*			10.988***			5.462**		
	Adjusted R^2	0.233			0.032			0.119			0.057		

Dependent variable: VRS efficiency score; significance: ****p*<0.001, ***p*<0.05; * *p*<0.1 (1-tailed)

Note: with regards to a regression equation with adjusted R^2less than 0.1, the low explanatory power suggests there are other unidentified exogenous factors related to the efficiency.

Table 4.29 Full sample: regression results for all types of efficiency in 2004-2006

		Cost efficiency			Academic efficiency			Academic effectiveness			Overall efficiency		
		Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.
Decile 1-3 group N=191	Intercept	0.737		0***	0.607		0***	0.787		0***	0.31		0***
	Ownership	0.12	0.409	0***	0.164	0.327	0***	0.028	0.063	0.514	0.286	0.386	0***
	Co-educated	-0.015	-0.052	0.496	-0.066	-0.13	0.104	-0.063	-0.141	0.114	-0.067	-0.089	0.265
	Mäori30%	0.021	0.093	0.183	0.046	0.121	0.097	0.015	0.045	0.577	0.051	0.089	0.222
	Std. Roll	0.485	0.562	0***	0.722	0.491	0***	0.291	0.225	0.008**	0.984	0.451	0***
	F-value	20.313***			15.131***			2.92**			14.488***		
	Adjusted R^2	0.289			0.229			0.039			0.221		
Decile 4-6 group N=316	Intercept	0.736		0***	0.701		0***	0.819		0***	0.503		0***
	Ownership	0.05	0.159	0.001**	0.119	0.313	0***	0	0.001	0.993	0.18	0.276	0***
	Co-educated	-0.035	-0.146	0.002**	-0.04	-0.141	0.006**	-0.069	-0.235	0***	-0.128	-0.261	0***
	Mäori30%	-0.003	-0.01	0.822	-0.069	-0.219	0***	-0.003	-0.01	0.856	-0.021	-0.039	0.442
	Std. Roll	0.398	0.614	0***	0.225	0.289	0***	0.159	0.201	0.001**	0.417	0.313	0***
	F-value	48.942***			27.095***			8.825***			24.632***		
	Adjusted R^2	0.378			0.249			0.09			0.231		
Decile 7-10 group N=349	Intercept	0.727		0***	0.78		0***	0.809		0***	0.573		0***
	Ownership	0.01	0.042	0.459	0.082	0.303	0**	0.01	0.047	0.475	0.08	0.175	0.008**
	Co-educated	-0.001	-0.005	0.922	-0.046	-0.182	0.002**	-0.047	-0.235	0***	-0.116	-0.269	0***
	Mäori30%	-0.104	-0.071	0.099*	-0.138	-0.084	0.092*	-0.117	-0.09	0.068*	-0.18	-0.064	0.19
	Std. Roll	0.357	0.631	0***	0.138	0.217	0***	0.186	0.37	0***	0.377	0.348	0***
	F-value	53.844***			17.032***			19.323***			21.2***		
	Adjusted R^2	0.378			0.156			0.174			0.188		

Dependent variable: VRS efficiency score; significance: ****p*<0.001, ***p*<0.05; * *p*<0.1 (1-tailed).

Note: with regards to a regression equation with adjusted R^2 less than 0.1, the low explanatory power suggests there are other unidentified exogenous factors related to the efficiency.

Table 4.30 Pure secondary school sample: regression results for all types of efficiency in 2004-2006

		Cost efficiency			Academic efficiency			Academic effectiveness			Overall efficiency		
		Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.
Decile 1-3 group N=141	Intercept	0.804		0***	0.619		0***	0.899		0***	0.372		0***
	Ownership	0.091	0.236	0.008**	0.302	0.454	0***	-0.021	-0.036	0.733	0.428	0.426	0***
	Co-educated	-0.06	-0.196	0.021**	-0.058	-0.109	0.183	-0.122	-0.262	0.01**	-0.108	-0.135	0.112
	Mäori30%	0.004	0.02	0.793	0.011	0.029	0.696	-0.014	-0.043	0.638	0.012	0.021	0.79
	Std. Roll l	0.46	0.574	0***	0.75	0.538	0***	0.194	0.159	0.095*	1.013	0.483	0***
	F-value	18.018***			20.888			2.913**			17.554***		
	Adjusted R^2	0.327			0.362			0.052			0.321		
Decile 4-6 group N=232	Intercept	0.775		0***	0.14		0***	0.838		0***	0.546		0***
	Ownership	0.04	0.105	0.069*	-0.054	0.285	0***	0.032	0.068	0.319	0.248	0.313	0***
	Co-educated	-0.035	-0.166	0.003**	-0.056	-0.194	0.001**	-0.056	-0.208	0.002**	-0.164	-0.368	0***
	Mäori30%	-0.013	-0.056	0.309	0.249	-0.183	0.002**	-0.005	-0.017	0.801	-0.036	-0.073	0.187
	Std. Roll	0.324	0.563	0***	0.14	0.332	0***	0.103	0.142	0.039**	0.438	0.364	0***
	F-value	27.878***			17.831***			3.918**			28.143***		
	Adjusted R^2	0.318			0.226			0.048			0.32		
Decile 7-10 group N=238	Intercept	0.811		0***	0.789		0***	0.864		0***	0.623		0***
	Ownership	0.036	0.167	0.012**	0.096	0.266	0***	0.007	0.027	0.712	0.082	0.14	0.055*
	Co-educated	-0.024	-0.153	0.011**	-0.04	-0.154	0.021**	-0.042	-0.231	0.001**	-0.117	-0.274	0***
	Mäori30%	-0.164	-0.138	0.013**	-0.201	-0.102	0.098*	-0.046	-0.034	0.586	-0.208	-0.065	0.283
	Std. Roll	0.216	0.546	0***	0.123	0.187	0.007**	0.124	0.275	0***	0.307	0.286	0***
	F-value	25.915***			8.569***			7.956***			11.601***		
	Adjusted R^2	0.296			0.113			0.105			0.152		

Dependent variable: VRS efficiency score; significance: ***p<0.001, **p<0.05; * p<0.1 (1-tailed).

Note: with regards to a regression equation with adjusted R^2less than 0.1, the low explanatory power suggests there are other unidentified exogenous factors related to the efficiency.

Since the expenditure data for the full sample were split from a mixed amount for both intermediate and secondary education, the approach for expenditure splitting may cause distortion on the scores of cost efficiency and overall efficiency to a certain extent; on the other hand, the exclusion of expanded schools in the pure secondary school sample can result in sample bias, as expanded schools contribute one third of the population. In this circumstance, the discussion below focuses on the common features of the regression results for both samples in each test period, while considering the weaknesses mentioned above.

It appears that for both test periods, a positive change in overall efficiency is significantly associated with integrated schools at all decile levels, consistent with the expectation that integrated schools outperform non-integrated schools in terms of maximizing academic outcomes using limited school resources. The decomposition results of overall efficiency show that for both test periods, cost efficiency and academic efficiency are positively related to integrated schools in decile 1-6, and also, there is no evidence showing that academic effectiveness is associated with integrated schools in decile 1-3 and decile 4-6 groups. The major difference in the regression results for integrated schools between the two test periods is associated with the decile 7-10 group: during the later period, academic effectiveness is positively related to them whereas during the earlier period, academic efficiency is positively associated with them. This indicates that between the two test periods, integrated schools might on average have an improvement in their capability of achieving entrances for tertiary education, given this is significant in decile 7-10 group which contributes 65% of integrated schools to the sample. A further investigation for this is undertaken in section 4.6.2 to look at the relationship between productivity change and integrated schools.

Results for both test periods show that overall efficiency is negatively related to co-educational schools in decile 4-10, which contributes more than 70% of co-educational schools to the sample. This is in keeping with the finding from prior research that on average, school overall performance is significantly and negatively associated with co-educational schools. During both the test periods, academic efficiency is negatively correlated to co-educational schools in decile 4-10 and academic effectiveness is also negatively associated with at least two decile groups of co-educational schools (2009-2011: decile 1-3 and decile 7-10; 2004-2006: decile 4-6 and decile 7-10). The main difference in the results for co-educational schools between the two test periods is related to decile 1-3 group. In contrast with the earlier period, the negative relationships between efficiencies and co-education schools during the latter period are significant. A further review of the dataset revealed that decile 1-3 group in the latter period contains a higher percentage of co-educational schools compared to that in the earlier period, this is the likely cause for the significant relationship, consistent with the view that co-educational schools underperform single-sex schools in terms of academic achievements.

With regards to Mäori30% schools, results from both test periods show there is no evidence suggesting that any change in overall efficiency is significantly associated with Mäori30% schools. During the latter test period, academic effectiveness is negatively related to Mäori30% schools in decile 1-3 while during the earlier period, academic efficiency is negatively correlated to Mäori30% schools in decile 4-6 group, in keeping with the fact that Mäori30% schools underperform non-Mäori30% schools in terms of academic achievements. Overall, between the two periods, the increases in academic outcomes (i.e., UE passes and NCEA L3 M&E) are less than the increases in academic outputs (i.e., NCEA qualification passes) in Mäori30% schools. Therefore, for Mäori30% schools, there is significant scope for improvement in achieving UE and NCEA endorsements yet to be made.

Results for both test periods confirm that cost efficiency and overall efficiency are positively associated with school roll at all decile levels, in keeping with the expectation that large schools are more efficient in terms of best utilizing school resources. The differences in the results between the two test periods are, in the earlier period, both academic efficiency and academic effectiveness are positively correlated to school roll for schools at all decile levels; while in the latter period, the positive relationship between academic efficiency and roll is not significant for decile 7-10 schools, as a result of some large decile ten schools[89] running CIE in parallel with the NCEA system. Moreover, the relationship between academic effectiveness and roll is not significant for decile 1-3 group, as a consequence of the improvement in NCEA qualification passes in small decile 1-3 schools, in particular, Mäori30% schools. This will be investigated further in section 4.6.2.

[89] Examples for these large schools are Macleans College, Auckland Gramma School and Westlake College.

4.6.2 Regressions on productivity change

A similar OLS regression equation was also applied to examine the relationship between productivity change and school characteristics of interest:

Equation 4.2 **Regression equation for productivity change**

$$I_i^t = \alpha + \beta_1 O^t + \beta_2 C^t + \beta_3 R^t + \beta_4 M^t + \beta_5 D^t + \varepsilon$$

Where, *I* is the productivity change from 2004 to 2011, measured by relative efficiency change/frontier-shift/Malmquist index (m0) under the categorical VRS model; *i* represents any one of cost efficiency, academic efficiency, academic effectiveness and overall efficiency; *t* signifies year 2004; *O* is the ownership indicator, where the value 1 means integrated and 0 means non-integrated; *C* is the co-educated school indicator, where the value 1 refers to co-educated schools and 0 indicates single-sex schools; *R* is standardized school roll; *M* is the indicator of Mäori density, where the value 1 represents schools with at least 30% Mäori students and 0 refers to schools with less than 30% Mäori roll; *D* reflects school decile level, where the value 1 indicates decile 1-6 whereas 0 indicates decile 7-10.

In this study, Malmquist index and its decomposed parts reflect school productivity change between 2004 and 2011. For this reason, only those schools with data in both 2004 and 2011 were included in the sample. As a consequence, 277 schools remained for the full sample and 199 schools were included for the pure secondary school sample.

Given that Mäori30% indicator was included in the regression model, decile level indicator was also used as a dummy variable to represent two groups: decile 1-6 and decile 7-10. The reason for not calculating the regression for the three decile groups as in the last section was, there are only 199 schools in the pure secondary schools sample and the sample size does not satisfy the minimum sample size[90] requirement for a regression analysis. Regression results are detailed in Tables 4.31-4.32 for the full sample and the pure secondary school sample respectively.

Since the approach for expenditure splitting only influences cost efficiency and overall efficiency, whereas the pure secondary school sample excludes expanded schools, discussions related to academic efficiency and academic effectiveness will focus on the

[90] According to Coakes (2012, p.140), the minimum sample size should be 20 times the number of independent variables in the regression model. Therefore, for this regression, the minimum size is 100 and running regressions by three decile groups does not meet this requirement.

regression results for the full sample, discussions associated with other efficiencies will incorporate the results for both samples.

Results in Tables 4.31 and 4.32 show that during the test period (i.e. 2004-2011), a negative productivity change in cost efficiency is significantly associated with Mäori30% schools in the full sample due to the significant backward frontier-shift. The backward frontier-shift is also significant in the pure secondary school sample but the negative relationship between the overall productivity change and Mäori30% schools is not significant. This suggests that since the introduction of NCEA, Mäori30% schools on average suffered a backward frontier-shift for cost efficiency in terms of best utilizing school resources, given the increased government grants they received from 2004 to 2011.

A positive overall productivity change in academic efficiency is significantly associated with Mäori30% schools and decile 1-6 schools. The positive overall productivity change for Maor30% schools is attributed to the improvement in relative efficiency and that for decile 1-6 schools is owing to the forward frontier-shift. As such, Mäori30% schools, low and medium decile schools had significant improvement in achieving NCEA qualification passes since the introduction of NCEA, consistent with the discussion related to Mäori30% schools and low decile schools in section 4.6.1.

It appears that since the introduction of NCEA, a positive overall productivity change in academic effectiveness is significantly related to integrated schools due to the significant forward frontier-shift, in keeping with the discussion about integrated schools in section 4.6.1, as well as the finding in section 4.5.4 that many expanded schools enjoyed a positive productivity change in academic effectiveness due to the forward frontier-shift since the introduction of NCEA, given that many expanded schools are integrated schools.

Results for both samples suggest that a positive overall productivity change in overall efficiency is significantly associated with co-educational schools, even though the significance for the pure secondary school sample is weaker compared to the full sample. Results for the full sample show that a negative overall productivity change in overall efficiency is significantly related to school roll due to the decrease in managerial efficiency, but these negative relationships are not significant for the pure secondary school sample. In short, since the introduction of NCEA, a positive overall productivity change in overall efficiency is more associated with co-educational schools than single-sex schools, while a negative overall productivity change in overall efficiency is more related to larger schools compared to smaller schools.

In summary, since the introduction of NCEA, integrated schools enjoyed an overall increase in productivity for academic effectiveness; co-educational schools had a positive overall productivity change in overall efficiency; Mäori30% schools suffered a backward frontier-shift for cost efficiency but enjoyed an overall improvement in productivity for academic efficiency; large schools suffered an overall decline in productivity for overall efficiency; low and medium decile schools enjoyed an overall increase in productivity for academic efficiency. The findings of positive productivity changes related to Mäori30% schools, decile 1-6 schools and co-educational schools provide supportive evidence for the achievement of the MoE’s strategic objectives, stressing “equity and excellence in education”.

Table 4.31 Full sample: regression results for productivity change

		Efficiency change			Frontier shift			m0		
		Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.
Cost efficiency	Intercept	0.917		0***	0.972		0***	0.891		0***
	Ownership	0.004	0.016	0.825	0.018	0.096	0.146	0.023	0.081	0.267
	Ced	0.023	0.103	0.125	0.007	0.041	0.494	0.03	0.121	0.066
	Mäori30%	-0.021	-0.078	0.245	-0.05	-0.263	0***	-0.063	-0.215	0.001**
	Std Roll	0.02	0.034	0.618	-0.04	-0.095	0.123	-0.021	-0.032	0.638
	Decile 1-6	0.031	0.147	0.031**	-0.046	-0.304	0***	-0.017	-0.073	0.269
	F-value	1.731			17.074***			4.292**		
	Adj. R^2	0.013			0.226			0.056		
Academic efficiency	Intercept	1.046		0***	1.029		0***	1.073		0***
	Ownership	-0.039	-0.076	0.296	0.054	0.129	0.062	0.016	0.021	0.768
	Ced	0.031	0.07	0.284	0.01	0.027	0.662	0.044	0.065	0.296
	Mäori30%	0.082	0.157	0.018**	0.045	0.104	0.094	0.161	0.202	0.001**
	Std Roll	-0.1	-0.087	0.201	-0.002	-0.002	0.973	-0.097	-0.055	0.396
	Decile 1-6	0.021	0.049	0.459	0.122	0.358	0***	0.146	0.228	0***
	F-value	4.032**			10.71***			9.812***		
	Adj. R^2	0.052			0.15			0.138		
Academic effectiveness	Intercept	1.025		0***	1.024		0***	1.044		0***
	Ownership	0.035	0.052	0.488	0.072	0.218	0.001**	0.116	0.153	0.04**
	Ced	0.03	0.052	0.436	-0.043	-0.152	0.012**	-0.019	-0.029	0.662
	Mäori30%	-0.036	-0.053	0.435	-0.048	-0.142	0.019**	-0.079	-0.102	0.127
	Std Roll	-0.043	-0.028	0.685	0.122	0.164	0.009**	0.106	0.062	0.368
	Decile 1-6	0.08	0.146	0.033**	-0.039	-0.145	0.019**	0.04	0.064	0.344
	F-value	1.339			15.008***			2.087*		
	Adj. R^2	0.006			0.202			0.019		
Overall efficiency	Intercept	1.213		0***	1.24		0***	1.494		0***
	Ownership	-0.063	-0.031	0.666	0.185	0.276	0***	0.134	0.046	0.525
	Ced	0.19	0.11	0.094	0.085	0.147	0.026**	0.362	0.144	0.029**
	Mäori30%	-0.18	-0.089	0.178	0.062	0.091	0.168	-0.096	-0.032	0.623
	Std Roll	-0.87	-0.193	0.005**	-0.042	-0.028	0.682	-1.201	-0.183	0.007**
	Decile 1-6	0.204	0.125	0.06**	-0.01	-0.019	0.772	0.244	0.102	0.124
	F-value	4.084**			4.371**			4.325**		
	Adj. R^2	0.053			0.058			0.057		

Dependent variable: VRS efficiency score; significance*: ***p*<0.001*, **p*<0.05*; * p*<0.1. (1-tailed).

Note: with regards to a regression equation with adjusted R^2 less than 0.1, the low explanatory power suggests there are other unidentified exogenous factors related to the efficiency or technical change.

Table 4.32 Pure secondary school sample: regression results for productivity change

		Efficiency change			Frontier shift			m0		
		Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.	Beta	Std. Beta	Sig.
Cost efficiency	Intercept	0.936		0***	0.92		0***	0.862		0***
	Ownership	-0.035	-0.109	0.175	0.013	0.072	0.371	-0.02	-0.061	0.446
	Ced	0.019	0.092	0.222	0	-0.001	0.988	0.018	0.085	0.258
	Mäori30%	0.004	0.016	0.836	-0.023	-0.165	0.037**	-0.016	-0.065	0.408
	Std Roll	0.036	0.069	0.387	0.015	0.05	0.538	0.046	0.086	0.286
	Decile 1-6	0.014	0.071	0.377	0.006	0.05	0.538	0.017	0.087	0.278
	F-value	1.748			1.295			1.32		
	Adj. R^2	0.019			0.007			0.008		
Academic efficiency	Intercept	1.12		0***	1.005		0***	1.124		0***
	Ownership	-0.062	-0.096	0.23	-0.025	-0.092	0.17	-0.087	-0.116	0.127
	Ced	0.004	0.01	0.894	0.023	0.13	0.038**	0.024	0.049	0.485
	Mäori30%	0.056	0.119	0.13	0.021	0.1	0.124	0.086	0.155	0.036**
	Std Roll	-0.136	-0.13	0.105	-0.107	-0.237	0***	-0.228	-0.187	0.014**
	Decile 1-6	0.003	0.007	0.926	0.062	0.367	0***	0.067	0.149	0.05**
	F-value	1.801			19.54***			6.614***		
	Adj. R^2	0.02			0.319			0.124		
Academic effectiveness	Intercept	0.983		0***	1.016		0***	0.996		0***
	Ownership	-0.019	-0.021	0.791	0.065	0.125	0.085	0.044	0.042	0.603
	Ced	0.045	0.079	0.294	-0.033	-0.099	0.141	0.01	0.015	0.841
	Mäori30%	-0.04	-0.061	0.437	-0.061	-0.157	0.026**	-0.09	-0.115	0.145
	Std Roll	-0.001	0	0.995	0.196	0.232	0.002**	0.21	0.122	0.132
	Decile 1-6	0.085	0.159	0.049 **	-0.064	-0.203	0.005**	0.015	0.024	0.768
	F-value	1.303			11.105***			1.223		
	Adj. R^2	0.008			0.203			0.006		
Overall efficiency	Intercept	1.146		0***	1.113		0***	1.301		0***
	Ownership	0.002	0.001	0.991	0.102	0.095	0.232	0.088	0.033	0.678
	Ced	0.123	0.11	0.146	0.071	0.103	0.165	0.221	0.132	0.081*
	Mäori30%	-0.184	-0.142	0.072	0.099	0.125	0.108	-0.039	-0.02	0.798
	Std Roll	-0.253	-0.089	0.272	-0.038	-0.022	0.786	-0.385	-0.09	0.263
	Decile 1-6	0.053	0.05	0.532	0.081	0.125	0.114	0.11	0.069	0.388
	F-value	1.247			2.616**			1.355		
	Adj. R^2	0.006			0.039			0.009		

Dependent variable: VRS efficiency score; significance: ****p*<0.001, ***p*<0.05; * *p*<0.1 (1-tailed).

Note: with regards to a regression equation with adjusted R^2less than 0.1, the low explanatory power suggests there are other unidentified exogenous factors related to the efficiency or technical change.

4.7 Comparison between top ten and bottom ten schools

This section concerns the third research question, which is "what are the possible causes for the performance of the top ten and bottom ten schools based on their published ERO reports". In order to answer this question, three steps were taken. First, the top ten and bottom ten schools were identified following a number of rules; second, statistics of causes for good/poor performance were produced based on the ERO reports of the selected schools; finally, a comparison was made between the top ten and bottom ten schools in terms of school characteristics and reasons for good/poor performance.

4.7.1 Selection of top ten and bottom ten schools

The selection of the top ten schools was based on the following:

- Each top ten school must be 100% efficient for at least three types of efficiency (cost efficiency, academic efficiency, academic effectiveness or overall efficiency);
- Each top ten school must have at least two school years sitting in the efficient group for each type of efficiency. This rule is to avoid the possibility that one school year occasionally receives 100% efficiency, given an ERO report reviews school performance for a period of several years;[91]
- ERO reports are available for the related test period (i.e., 2004-2006 or 2009-2011).

The selection of the bottom ten schools was based on the following:

- Each bottom ten school must have at least two school years sitting in the lowest quartile of all types of efficiency. This rule is to avoid the possibility that one school year occasionally receives a poor efficiency score whereas other school years for the same school have much better efficiency;
- Each bottom ten school must have the lowest overall efficiency. This rule is to determine which school is selected where there are multiple choices;
- ERO reports are available for the corresponding test period (i.e., 2004-2006 or 2009-2011).

Following the above rules, the top ten and bottom ten schools for the full sample were selected. Their major characteristics are listed in Appendix XV and XVI. Since the statistics for the pure secondary school sample are similar to the full sample with

[91] Generally, the interval for ERO review is 2-5 years, on average 3 years.

respect to the main features of the top ten and bottom ten schools, the following discussion mainly focuses on the results for the full sample.

The major characteristics of these schools are summarized as below:

For the period of 2009-2011, the top ten school group contains either large (roll>1,000) or extremely small (roll<200) schools, highest decile (decile 9-10) or lowest decile (decile 1-3) schools, consistent with the view that large and high decile schools tend to be more efficient due to economies of scale, small and low decile schools may get high efficiency scores in terms of academic efficiency, academic effectiveness and overall efficiency, as a result of the Government's assistance to low decile schools as well as the impact from the categorical VRS model, which favours extremely small and lowest decile schools as a result of adjusting the effect of size and the SES decile. There are four out of ten integrated schools in the group and five out of ten are girl schools. Of note, there are no boy schools selected in the top ten group, consistent with the fact that boy schools appear to underperform under the NCEA qualification system. There is one Mäori30% school in the top ten group, six out of ten schools have less than 10% of Mäori students. The average Mäori density for the top ten schools is 18.5%, similar to New Zealand school demographics (an average of 19.6% for the period of 2009-2011).

Statistics for the top ten schools during 2004-2006 are very similar to that in 2009-2011, Mäori density for the top ten schools is only 8.7%. Of note, there are six schools sitting in the top ten groups for both test periods.

With regard to the bottom ten schools, in 2009-2011, relatively small (300<roll<700), low and medium decile (decile 2-8) schools are found in this group; that is to say, large schools, decile one and decile 9-10 schools are not in the bottom ten group, consistent with the view that large schools, decile one and decile 9-10 schools have better capability of fundraising, and consequently can better engage student learning. This is because decile one schools get more financial help from the Government and decile 9-10 schools receive more donations from parents, leaving those medium decile schools with limited capacity to get additional financial support.

During 2009-2011, there is only one integrated school and one boy school identified in the bottom ten group, whilst there are no girl schools in this group, consistent with the previous findings that most integrated schools and girl schools outperform other schools under the NCEA qualification system. In contrast to the top ten group, there are five out of ten Mäori30% schools in the bottom ten group, and two of them have special educational objectives that focus on the needs of specific groups (e.g., Teen Parent Unit, Special Needs Unit). They are all decile 2-3 schools with rolls between 300 and 500. The average Mäori density in this group is 38%.

Results for the bottom ten schools in 2004-2006 are very similar to those in 2009-2011, with school rolls between 200 and 800, Mäori density is on average 30%. The only difference is, there is a decile one school in the bottom ten group.

4.7.2 Possible causes for school performance

To identify the likely causes for the performance of the top ten and bottom ten schools, 40 ERO reports of these schools in both test periods were reviewed. A number of performance indicators were recorded and interpreted into short phrases based on the literal meaning and the published description of school evaluation indicators (Education Review Office, 2011), examples for good practice[92] in these indicators are detailed in Table 4.33.

Drawing on the performance indicators, the frequency statistics (in percentage) of the likely causes for good and poor performance in the two test periods are illustrated in Figures 4.1-4.4.

For the 2009-2011 period, the top five reasons for good performance are good practice in school governance and managing, successfully engaging student learning, satisfying students' needs, practical teaching strategy and effective use of data. The top five causes for poor performance include problematic governance and managing, poor use of data, unsuccessfully engaging student learning, poor practice in teaching strategy and ineffective curriculum delivery. Aside from the quality of practice, the only difference in the causes for performance between the top ten and bottom ten schools in this period is, the top ten schools can satisfy students' needs whereas the bottom ten schools cannot deliver reasonable curriculum. Potentially, these two causes have an association as the bottom ten schools cannot identify and understand students' needs; as a result, they cannot design/delivery appropriate curriculum to satisfy students' needs.

For 2004-2006 period, the top five causes for good performance include successful teaching strategies, operative student support, sound practice in governance and

[92] Examples for poor practice in these indicators are not included in Table 4.33 because they are simply the reverse of the good practices.

managing, comfortable learning environment and effective use of data. The top five causes for poor performance are problematic practice in governance and managing, defective teaching strategy, substandard assessment strategies, being unable to satisfy students' needs and poor use of data. The different causes for the performance between the top ten and the bottom ten schools suggest that the top ten schools understand students' needs therefore they can provide effective student support and good learning environment to engage student learning. In contrast, the bottom ten schools have poor practice in understanding students' needs and analysing students' assessment data therefore they are unable to set suitable targets for students to improve their academic performance. As a consequence, they cannot ensure an effective assessment system in place to encourage student learning.

It appears that in both test periods, good performance mainly depends on a sound governance and managing practice. However, besides this, it seems that good practice in the later period relied more on engaging student learning and use of data whereas in the earlier period, good practice emphasized teaching strategy and student support. One reason for these differences is that at the time of the introduction of NCEA, teaching strategies and student support policies in some schools needed to be developed or adjusted in line with the new qualification system. Another reason is in the later period, data analysis was getting more and more important for schools to develop assessment strategies, design curriculum and introduce new programmes to engage student learning. The last reason is, since 2007, the Ministry and the ERO expect schools to retain as many students until year 13 as possible. Therefore, engaging student learning is the primary educational objective for each school in the later test period.

Table 4.33 Examples of indicators for good school performance

Performance indicators	Examples of indicators
Assessment strategies	Valid and reliable information collected from different sources; fair and inclusive assessment process;
Curriculum delivery	Link to school's goals and national standard: *The New Zealand Curriculum;* clear and coherent content; reflect student needs and community desires; enhance student engagement and achievement;
Engaging student	High level attendance and retention rate; high levels of student interest and motivation in learning;
Family & community	Engage parents and community with diverse needs for student learning;
Financial management	Effective financial planning and budgeting;
Governance &managing	Clear school's vison and values; effective strategic planning for student achievement; appropriate allocation of school resources; high quality decision-making; sound performance management system for school principals and teachers;
Health and safety	Effective policies and procedures in place to address any concerns like bullying and harassment and also promote health and safety for students, staff and visitors;
Learning environment	Positive learning, safe and emotional environment;
Compliance	Comply with financial, property, human resource, health and safety and other legal requirements;
School culture	Respectful relationships; diverse students with different background;
Self-review	Effective self-review process: gather evidence relating to student progress and achievement, teaching practice and resource utilizing; use the evidence to monitor performance towards targets; plan to improve;
Student support	Effective policies and procedures to support student learning and achievement
Students' needs	Teaching strategies and curriculum design match gifted and talented, and special needs via feedback from students and parents, assessment data analysis and class observation;
Teaching strategy	Choose teaching strategies that best meet students' needs and ability; deliberate acts of teaching to engage students in purposeful learning;
Use of data	Gather and analyse valid and reliable information about students such as assessment results and attendance data; use the analysis results to plan for improvement in teaching practice or school governance and management.

Extracted from Education Review Office (2011)

Figure 4.1 Causes for good performance in the period 2009-2011

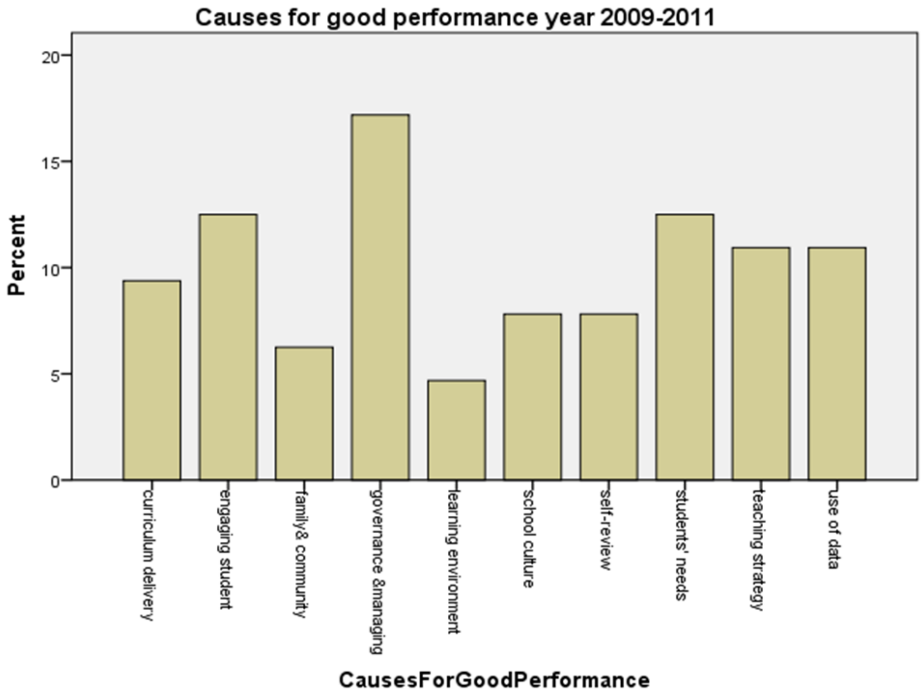


Figure 4.2 Causes for poor performance in the period 2009-2011

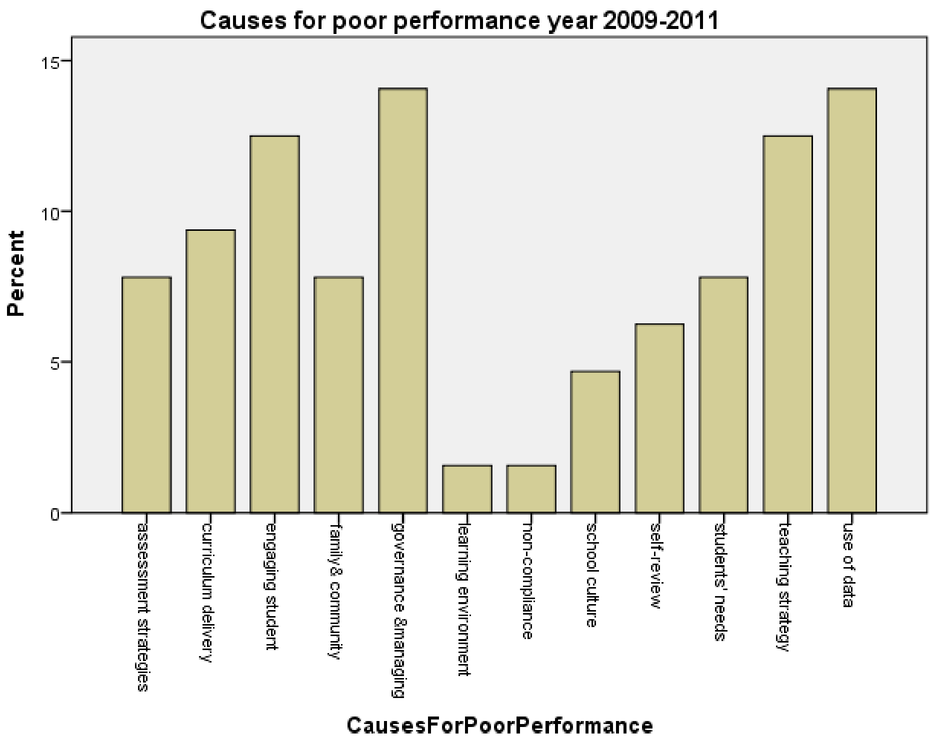


Figure 4.3 Causes for good performance in the period 2004-2006

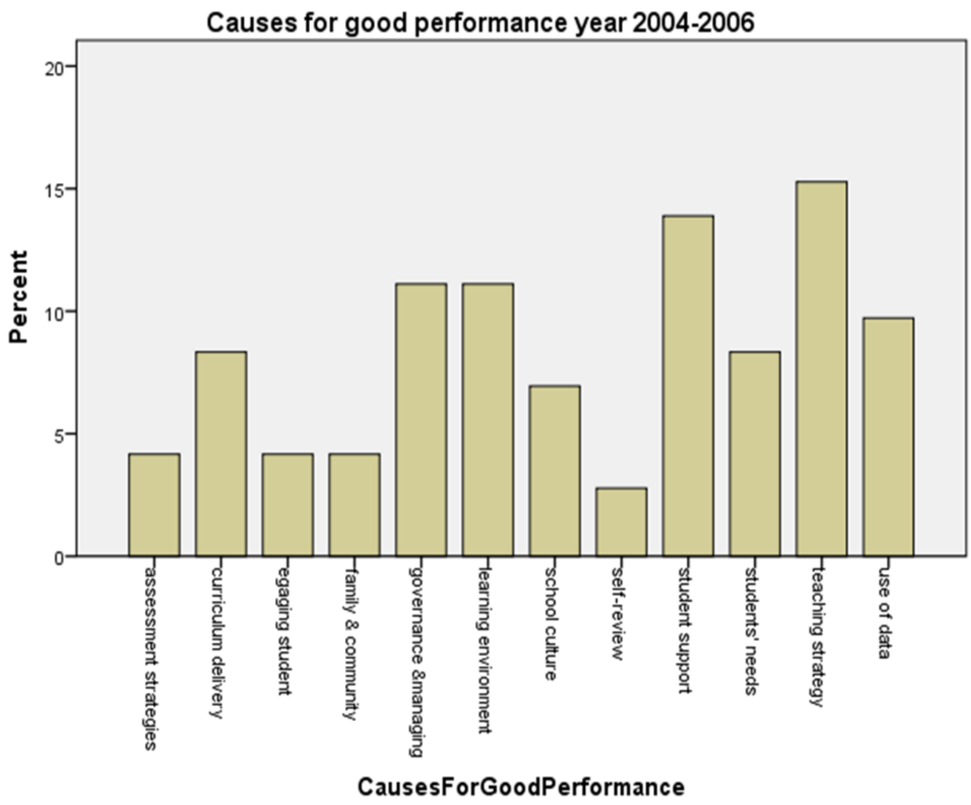
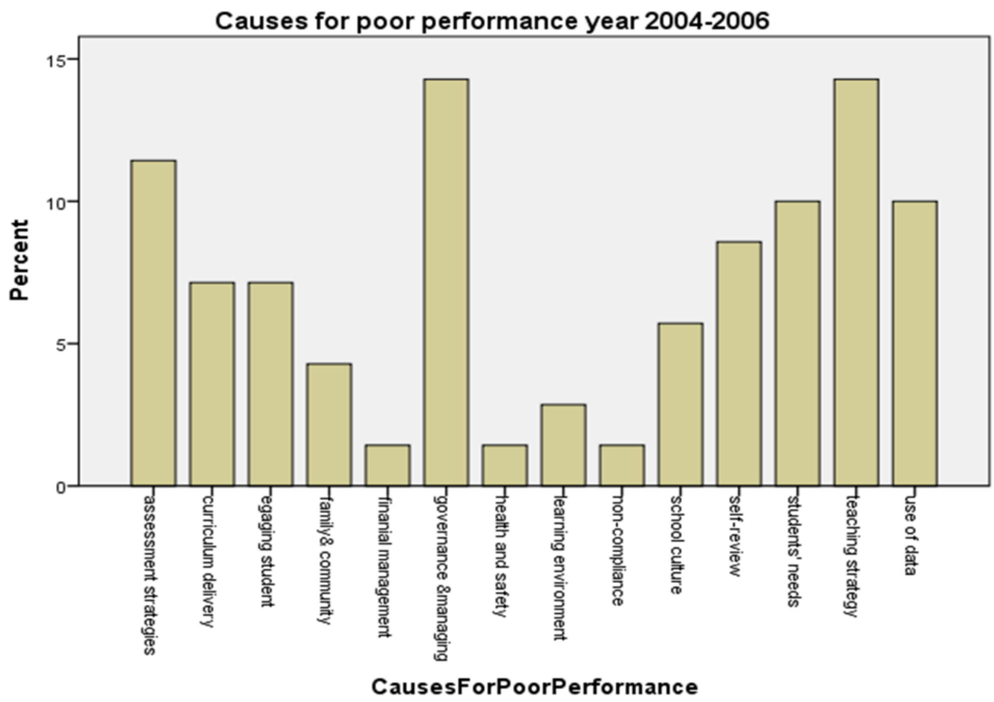


Figure 4.4 Causes for poor performance in the period 2004-2006



4.8 Summary and discussion

This chapter investigated the performance of New Zealand public secondary schools under the NCEA system based on two samples: the full sample and the pure secondary school sample. As the pure secondary school sample is considered to be more homogeneous in terms of the educational process, data analysis results for this sample are likely to be more robust than that for the full sample.

However, the results for the pure secondary school sample might not adequately reflect the characteristics of New Zealand secondary school performance due to the exclusion of expanded schools, which contribute about one third to the target population. In fact, the results received for both samples are generally consistent with respect to the direction of association and level of significance for most analyses undertaken. As such, the full sample is considered more representative of the underlying secondary school population in New Zealand. Accordingly, all the discussions in this paper are primarily based on the results for the full sample.

Based on the framework developed in section 2.3, the output-oriented categorical VRS model was applied to all the efficiency models in the framework: cost efficiency, academic efficiency, academic effectiveness and overall efficiency models, given that not only can the categorical VRS model adjust for the difference in school size but also provide best control for the effect of student socio-economic backgrounds. Results show that large and high decile schools generally outperform small and low decile schools in terms of all the efficiencies, despite some large decile ten schools that mainly focused on the CIE system rather than the NCEA system during the test periods.

The differences between overall efficiency and its sub-efficiencies demonstrate that the network DEA model can better reflect school performance from multiple facets, compared to the overall efficiency model which only provides a single score that hides some crucial information for performance evaluation and improvement. In this sense, the application of a network DEA model in this study extended the previous framework for public school performance and hence adds to the literature by modelling the performance measurement system through a decomposition of the overall efficiency into a series of sub-efficiencies.

Relevant regression results show that school characteristics are associated with school performance efficiencies. The findings outlined as follows are consistent with those from prior research: with regards to academic performance, integrated schools outperform non-integrated schools; co-educational schools underperform single-sex schools and Mäori students lag behind non-Mäori students though the difference is gradually reduced.

The results of the Malmquist index and its decomposition for all the efficiencies show that since the introduction of NCEA, school productivity overall increased in terms of academic efficiency, academic effectiveness and overall efficiency. In particular, forward frontier-shifts were identified for both academic efficiency and overall efficiency, whilst the positive productivity change in academic effectiveness was due to the increase in relative efficiency. However, a negative productivity change in cost efficiency was also identified, as a result of the significant increase in teaching expenditures across the test period.

A further examination of the relationships between productivity change and school characteristics revealed that since the introduction of NCEA, integrated schools enjoyed an overall increase in productivity for academic effectiveness; co-educational schools had an overall positive productivity change in overall efficiency, even though they are still less efficient than single-sex schools; Mäori30% schools suffered a backward frontier-shift for cost efficiency but had an overall improvement in productivity for academic efficiency; low and medium decile schools had an overall increase in productivity for academic efficiency. These findings suggest that NCEA levels the playing field for different student groups categorized by school ownership, gender, school operating scale and SES deciles. Furthermore, regression results also suggest that large schools had an overall decline in the productivity for overall efficiency, even though they are considered to be more overall efficient than small schools; likewise, single-sex schools suffered a decrease in productivity for overall efficiency whereas they had better overall efficiency on average than co-educational schools. Therefore, it is concluded that NCEA has narrowed the gap of performance between different student groups, given that the negative productivity change during the test period are more associated with the more efficient groups (i.e., large schools, single-sex schools) than their counter-parties.

The examination of the possible causes for good/poor performance revealed that a good school performance primarily depends on sound governance and managing practice. During the later test period, good practice relied more on engaging student learning and use of data whereas in the earlier period, good practice emphasized teaching strategy and student support. The difference in good school practice reflects the evolutionary process of the NCEA system to a certain extent,[93] along with the change in the MoE's strategic objectives.

[93] Since 2007, certificates for NCEA Merit and Excellence endorsements are issued to get hardworking students recognized.

Chapter 5 Conclusion

5.1 Introduction

This chapter discusses the conclusion on the research objectives and key features of school performance under the NCEA system; some recommendations are made to policy reformers. Moreover, the main contributions of this study are identified, followed by the description of research limitations and suggestions for future research.

As mentioned in chapter 1, the first research objective was to develop a model of performance measurement system for New Zealand public secondary schools under the NCEA qualification. Based on a number of theoretical perspectives drawn from management accounting, management control and educational economics, this model incorporated performance measures that can reflect the MoE's objectives, which are considered the expectations of the primary stakeholder group in the New Zealand educational context.

Given that public school performance should reflect the efficiency of utilizing resources and the effectiveness of achieving organizational objectives, the prior framework for school performance measurement was extended. In addition to measuring the efficiency for producing outputs from resources, the effectiveness for achieving outcomes from outputs was also considered.

The second research objective was to evaluate the NCEA qualification system by looking at whether there was an overall productivity change since the introduction of NCEA, given that it had been criticized for being less reliable, girl-friendly and for encouraging mediocrity (NZ Herald, 2007). The research findings reveal that there was a significant increase in productivity for overall efficiency since the introduction of NCEA due to the forward frontier-shift. In particular, the improvement in productivity for academic efficiency was significant due to the increase in relative efficiency and the forward frontier-shift, whilst the overall productivity change in academic effectiveness was also positive and significant owing to the increase in relative efficiency. The regression results also suggested that low and medium decile schools or high Mäori density schools had overall a significant improvement in productivity for academic efficiency, whilst co-educational schools on average had a significant increase in productivity for overall efficiency. On the other hand, regression results also indicate that the negative productivity change in overall efficiency is more related to large schools or single-sex schools compared to small schools or co-educational schools. Accordingly, these findings have confirmed that the NCEA system levels the playing field for different student groups. As a result, the gaps in performance between them are narrowed, in line with the expectations of the Government and the MoE, stressing 'equality and excellence in education".

The third research objective was initially to check whether the research results relating to schools with good or poor performance are consistent with the ERO evaluation. In this regard, the top ten and bottom ten schools were identified and possible reasons for good or poor performance were extracted from their ERO reports. The comparison results on the performance between the top ten and bottom ten schools show that school practice in 'governance and managing' is a double-edged sword, which is the cause for both good and poor performance of New Zealand public schools. Furthermore, good school practice during the later period (2009-2011) focused more on engaging student learning and use of data whereas in the earlier period (2004-2006), good practice related to teaching strategy and student support. These findings suggest that the change in the evaluation standards for school performance is in parallel with the evolution of the NCEA system, which focuses more on improving the reliability of internal assessment standards and encouraging students to achieve high level NCEA endorsements.

5.2 Key features of school performance under the NCEA system

Since some features of the New Zealand educational system remained unchanged after the introduction of NCEA, the following conclusions were drawn from the comparison with a prior study (i.e., Harrison (2008)) on New Zealand school performance under the old qualification system (i.e., School Certificate, University Entrance, Sixth Form Certificate and University Bursary qualifications).

Given that there has been no change in regards to the structure of the school accountability system, the homogeneity of education process, as well as the implication of SES decile since the introduction of NCEA, a performance measurement framework from primary stakeholder perspective was established to measure public school performance under the NCEA system by drawing on the work of Harrison (2008). Subsequently, an output-oriented categorical VRS DEA model was employed in this study to calculate school performance efficiencies from different dimensions. Therefore, the fundamentals of performance measurement method in this research is similar to that of Harrison (2008), in the sense that school performance is reflected by the productive efficiency of New Zealand education process, which is modelled by the DEA approach.

Since this study measures school performance using multiple efficiencies whereas Harrison (2008) calculates a single composite efficiency to examine school performance, model variables employed in these two studies are significantly different, the comparison of school performance under the NCEA system to the old qualification system can only be undertaken on a qualitative basis.

With regard to the decile effect, findings for NCEA and the old system are similar: higher decile schools outperform lower decile schools in terms of maximizing academic outputs; decile one schools have better academic performance than their adjacent decile schools as a result of the targeted assistance from the MoE. However, the academic performance of decile ten schools under the NCEA system is not as outstanding as that under the old system compared to the lower decile schools, because after the introduction of NCEA, some decile ten schools focused more on the CIE system and thus obtained fewer NCEA academic outputs than expected.

For both the NCEA and the old systems, school roll drives cost efficiency due to the economics of scale. That is, large schools tend to be more efficient in terms of resource utilization irrespective of what qualification system is operated.

From a technical view, for both qualification systems, many schools are operating at the most productive scale, but a higher proportion of schools are running under ‘increasing returns to scale’ (IRS). This implies that theoretically, the overall productive efficiency of New Zealand public schools could be improved by increasing school size.

With respect to the difference in performance between different groups categorized by school characteristics, findings for both the qualification systems are similar. Single-sex schools outperform co-educational schools in terms of maximizing academic outputs; integrated schools are more efficient than non-integrated schools in terms of academic achievements. A detailed examination for girl school performance in this study revealed that girl schools outperform both boy schools and co-educational schools under the NCEA system in terms of academic achievements, consistent with the finding in Alexander et.al (2010), showing that New Zealand girl schools have better academic performance under the old system, compared to boy and co-educational schools. Accordingly, the criticism of ‘NCEA favors girls’ is not convincing. In fact, gender difference in New Zealand is consistent with international data, for example, statistics for Programme for International Student Assessment (PISA) and General Certificate of Secondary Education (GCSE) demonstrate that the situation in other countries is the same as that in New Zealand (New Zealand Qulification Authority, 2016a).

Given the significant differences in both test period setting and efficiency measures for Malmquist analysis between this study and Harrison (2008), it is hard to compare school productivity changes between the two systems. However, it appears that under any of them, there was a significant decline in overall productivity change in terms of resource utilization due to the increased Government grants into schools across the test periods, consistent with the fact that the Government made continuous effort to improve the overall performance of public schools.

Nevertheless, as mentioned above, the key finding of this study is, there was a significant improvement in productivity for overall efficiency and its sub-efficiencies being academic efficiency and academic effectiveness. Among others, forward frontier-shifts were identified for both overall efficiency and academic efficiency. This confirms the success of the NCEA system in improving the overall productivity for New Zealand public schools in terms of maximizing academic outputs from the budgeted resources. While in contrast, for the old system in several years before 2001, there was no significant change or even a decline was identified in school productivity with respects to maximizing academic achievements or minimizing school costs according to both Harrison (2008) and Jaforullah (2010); though their model variables are different, both of them include academic results as outputs and school resources as inputs.

As discussed earlier, NCEA narrows the gap in performance between different student groups categorized by school characteristics and Mäori density. This is consistent with the initial intention for the introduction of NCEA, which equipped students for 'life-long learning', rather than was just targeted to selecting the elite for tertiary education.

With regard to the criticism of 'NCEA encourages mediocrity', the NZQA improved the qualification system by issuing certificates for NCEA Merit and Excellence endorsements since 2007, in order to encourage able students to achieve high levels of recognition. As a consequence of this action, a positive change in school academic effectiveness is expected during the period of 2004-2010; thus the result of Malmquist indices for academic effectiveness is consistent with the expectation. However, the increase in productivity for academic effectiveness is attributed to the improvement in relative efficiency, rather than the effect of policy change reflected by a forward frontier-shift. This suggests that additional steps from the MoE or the NZQA are needed to make further achievement (i.e., forward frontier), as research reveals that it is hard for a school to improve/retain its relative efficiency in the same sample over three continuous years (Thomas, Peng and Gray, 2007).

The improvement in productivity for academic efficiency of Mäori30% schools is of importance because supporting the Mäori group to achieve NCEA qualifications is a strategic goal of the MoE, given that the official statistics show that the academic performance of Mäori students lags behind other ethnic groups. However, the positive overall productivity change is owing to the increase in relative efficiency, rather than the forward frontier-shift. Given that this increasing trend might reverse after three years (Thomas et.al, 2007), additional policies are needed to assist in performance improvement for the Mäori group. For example, schools could be empowered to develop more curriculum or introduce more programmes to engage Mäori student learning.

Since ‘governance and managing’ was identified as the most likely reason for school success/failure, the MoE needs to formulate effective policies that help schools strengthen the practice in school governance and management. Moreover, the dissemination of good experience from the efficient peers is also encouraged in the learning network. The differences of possible causes for good/poor performance between the later and earlier periods reflect the dynamic nature of the education system, to which the NCEA system, the national curriculum standards, as well as school performance measurement/management systems need to be adapted.

5.3 Contributions

The main contributions of this thesis are described as below:

First, this research extended the prior framework of performance measurement for New Zealand public schools, which was developed by Harrison (2008). The extended framework decomposes the single performance measurement process into three inter-connected sub-elements that reflect school performance from multiple dimensions, being cost efficiency, academic efficiency and academic effectiveness. These sub-efficiency models provide more performance information that is hidden by the single overall efficiency model.

The extended framework adds to the literature on public school performance measurement by looking at both efficiency and effectiveness of school education process, given that the prior quantitative research only calculated productive efficiency to measure New Zealand public school performance. In this sense, the approach has provided an exemplar method of measuring organizational performance in the public sector, where both efficiency of resource utilization and effectiveness of organizational objective achievement are essential to performance measurement for public organizations.

To ensure that the proposed framework can properly reflect school performance under the NCEA system, a number of new performance measures were identified to represent school academic outcomes and outputs (Table 2.1). Of note, other than just incorporating the primary stakeholders' objectives, the proposed framework can be used to satisfy the expectations of multiple stakeholders through selecting performance measures in alignment with their objectives. For example, parents might be more interested in school academic outputs and outcomes, additional measures like the number of scholarship, and the number of NCEA Level 1 and Level 2 Merit and Excellence endorsements could be added to the framework.

Second, this thesis employed a network DEA model to implement the framework. The sub-efficiencies were produced by a series of inter-connected DEA models and the overall efficiency was also calculated for comparison. The application of a network DEA model breaks the education process into three sub-processes and produces the corresponding sub-efficiencies to measure their productivity. As such, the network DEA model explicitly demonstrates the transformation process and calculates sub-efficiencies to measure all of the intermediate processes. In doing this, inefficient schools are able to improve their overall performance by focusing on the most inefficient sub-processes, rather than facing a 'black box' as is common in a single standard DEA model (e.g., the overall efficiency model in this study). As such, this study adds to the literature of DEA application in performance measurement for the public sector.

Finally, given that there had been limited quantitative research looking at New Zealand public school performance under the NCEA system, this thesis addressed this gap by conducting a large-scale quantitative research based on two samples: the full sample approximates the target population and the pure secondary school sample inclusively represents New Zealand pure secondary schools. Moreover, the empirical analysis of school productivity change also adds to the limited research on the effectiveness of NCEA, given that most research focuses on school academic performance and the criticisms against NCEA, rather than examines the impact of NCEA on school productivity from the perspective of education economics. As the Government has invested large expenditures on implementing and optimizing NCEA, its impact on school productivity is particularly important in the eyes of the Government and the public.

5.4 Research limitations and directions for future research

Since a categorical DEA model was employed to calculate performance efficiencies, potentially, the efficiencies for lower decile groups are biased upward. Moreover, the network DEA model also biases upward sub-efficiencies for small schools, given that in the New Zealand context, school expenditures, number of NCEA passes, UE passes and number of NCEA endorsements are driven by school roll. In this sense, efficiency results for very small schools should be interpreted with caution. It would be useful to look at school performance in conjunction with their ERO reports, in particular, for those small schools with polarized efficiencies.

There are a number of directions for future research in the area of performance measurement for New Zealand secondary schools. First, it would be interesting to develop a school benchmarking system by incorporating school quality indicators, given that DEA technique is a linear-programing-based methodology that has been established as an effective tool for certain types of benchmarking in many large organizations (Zhu, 2003). Second, a DEA model can also be used to measure school accountability, given that NCEA is one of the most transparent systems in the world as many NCEA academic results are published online at school level. According to Wöbmann et al. (2007), school accountability measures ‘that make students, teachers, and schools more responsible for their actions can lead to improved student achievement’. As such, whenever the academic results for external exams and internal assessment, agreement rates for internal assessment moderation are available, school accountability can be modelled using a DEA approach. Third, given that the NCEA qualifications can be recognised overseas, for example, in Australia, NCEA Level 3 is recognised as equivalent to Senior Secondary Certificates of Education (New Zealand Qualification Authority, 2016c). The DEA technique can be applied to model school performance in different countries and the results can be used to compare school performance between New Zealand and other countries. Finally, it has been a great challenge to link school performance efficiency to school ERO reports due to their qualitative nature. Harrison (2008) conducted a content analysis on ERO reports and found a strong and positive relationship between the DEA efficiency and ERO index. This finding is indeed an impetus for further research. However, an improved format and content of ERO reports would assist this process.

5.5 Concluding comments

The research findings provide favourable evidence for the success of NCEA, in the sense of supporting the MoE and the NZQA to carry forward, develop and improve current curriculum standards and qualification/endorsement assessment standards. However, sustaining the credibility of NCEA requires continuing alignment with international best practice in assessment, given that NCEA has been criticized for encouraging mediocrity and for using less consistent assessment standards compared to CIE or the International Baccalaureate (IB) Diploma Programme.[94] Therefore, the development and improvement of NCEA needs to be "undertaken in tandem with developments in the New Zealand Curriculum and new knowledge about teaching, learning and assessment" (New Zealand Qualifications Authority, 2014, p. 6).

The DEA model developed in this study could be adapted for this ongoing research, which examines the differences in educational outcomes under different qualification systems, in particular, comparing NCEA outcomes to CIE achievements in the New Zealand educational context. The findings of this research could assist the MoE and the NZQA in decision-making for sustaining the credibility and international competitiveness of the NCEA system.

[94] In New Zealand, only a small number of schools have adopted IB Diploma Programme therefore it is not discussed in previous chapters.

References

Alam, M. (2006). Stakeholder Theory. In Z. Hoque, Methodological Issues in Accounting Research: Theories and Methods (pp. 207-222). London: Spiramus Press Limited.

Alexander, W. R., Haug, A. A., and Jaforullah, M. (2010). A two-stage double-bootstrap data envelopment analysis of efficiency differences of New Zealand secondary schools. Journal of Productivity Analysis, 34(2), 99-110.

Amado, C. A., Santos, S. P., & Marques, P. M. (2012). Integrating the Data Envelopment Analysis and the Balanced Scorecard approaches for enhanced performance assessment. *Omega*, 40, 390-403.

Association of Teen Parent Educators New Zealand. (2011). *Our Schools.* Retrieved from Teen Parent Schools in New Zealand: http://teenparentschools.org.nz/our-schools

Atkinson, A. A., & McCrindell, J. Q. (1997). Strategic performance measurement in government. *CMA*, 7 1 Q), 20-23.

Atkinson, A. A., Waterhouse, J. H., & Wells, R B. (1997). A stakeholder approach to strategic performance measurement. *Slogan Management Review, spring*, 25-37.

Banker, R. D., & Morey, R. C. (1986). The use of categorical variables in Data Envelopment Analysis. *Management Science (1986-1998), 32*(12), 1613.

Banxia Software Limited. (2010). *An Introduction to Frontier Analyst 4.0.* Retrieved from Banxia Software: http://www.banxia.com/pdf/fa/FrontierAnalystWorkbook.pdf

Barnum, D. T., and Gleason, J. M. (2006). Biases in technical efficiency scores caused by intra-input aggregation; mathematical analysis and a DEA application using simulated data. *Applied Economics*, 38, 1593-1603

Beer, S. (1972), *Brain of the Firm*, Allen Lane/The Penguin Press, London

Bendikson, L., Hattie, J., & Robinson, V. (2011). Identifying the comparative academic performance of secondary schools. *Journal of Educational Administration*, 49(4), 433-449.

Bititci, U.S., Carrie, A. and McDevitt, L. (1997), “Integrated performance measurement systems: an audit and development guide”, *The TQM Magazine*, Vol. 9 No. 1.

Bititci, U.S., Turner, T. and Begemann, C. (2000), “Dynamics of performance measurement systems”, *International Journal of Operations & Production Management*, Vol. 20, pp. 692-704.

Brignall, S. and Modell, S. (2000), “An institutional perspective on performance measurement and management in the ‘new public sector’”, *Management Accounting Research*, Vol. 11, pp. 281-306.

Brudan, A. (2010), “Rediscovering performance management: systems, learning and integration”, *Measuring Business Excellence*, Vol. 14 No. 1, pp. 109-23.

Bosker, R., & Scheerens, J. (1989). Issues in the interpretation of the results of school effectiveness research. *International Journal of Educational Research*, 13, 741-751.

Calton, J.M. (1993), “What is at stake in the stakeholder model?”, in Ludwig, D.C. (Ed.), *Business and Society in a Changing World Order*, The Edwin Mellen Press, Lewiston, pp. 101-27.

Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 50(6), 1393-1414.

Charnes, A., Cooper, W. W. and Rhodes, E (1978) Measuring the efficiency of decision making units. *Enropean Joumal of Operational Retearch*, 2(6), 429-444.

Clarkson, M. B. E. (1995). A risk-based model of stakeholder theory. Paper presented at the 13th Annual Meeting of the Society for Business Ethics, Vancouver, Canada.

Coakes, S.J. (2012). *SPSS Version 20 for Windows: Analysis without Anguish*. John Wiley & Sons, Brisbane.

Coe, R., & Fitz-Gibbon, C. T. (1998). School effectiveness research: Criticisms and recommendations . *Oxford Review af Education*, 24(4), 421-438.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). New York: Springer.

Cook, W. D., Tone, K., & Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega*, 44, 1-4.

Crotty, M. (1998). The Foundations of Social Research: Meaning and perspective on the research process. St Leonards, NSW: Allen & Unwin. Chapter 1 Introduction: The research process, pp1-17.

Cross, K, and Lynch, R. (1989). Accounting for competitive performance. *Journal of Cost Management, Spring*, 20-28.

Debreu, G. (1951). The coefficient of resource utilization. *Econometric*, 19(3), 273-292.

Dyson, R., Allen, R., Camanho, A., Podinovski, V., Sarrico, C., and E., Shale (2001). Pitfalls and Protocols in DEA. *European Journal of Operations Research 132*, 245-259.

Education Counts. (2014b). *Directories.* Retrieved 2014, from Education Counts: http://www.educationcounts.govt.nz/directories/list-of-nz-schools

Education Review Office. (2011). *Evaluation Indicators for School Reviews 2011.* Wellington: New Zealand Government.

Education Review Office. (2014a). *What ERO Does.* Retrieved 2014, from Education Review Office: http://www.ero.govt.nz/About-Us/What-ERO-Does

Education Review Office. (2014c). *Framework for School Reviews 2014.* Wellington: New Zealand Government.

Education Review Office. (2014b). Statement of Intent 2013–16. Wellington: New Zealand Government.

Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1989, 1994). Productivity change in Swedish hospitals: A Malmquist output index approach. In A. Charnes, W. W. Cooper, A. Y. Lewin & M. L. Seiford (Eds.), *Data Envelopment Analysis: Theory, Methodology and Applications.* Boston: Kluwer Academic Publishers.

Färe, R., Grosskopf, S. (2000). Network DEA. *Socio-Economic Planning Sciences*, 34, 35–49.

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120(3), 253-281.

Feinstein, L., & Symons, J. (1999). Attainment in secondary school. *Oxford Economic Papers*, 51, 300-321.

Field, A. Miles, J. (2013). *Discovering Statistics using IBM SPSS Statistics*, Fourth Edition. Sage, Thousand Oaks.

Fiske EB, & Ladd H (2000). When school compete: a cautionary tale. Brookings Institute Press, Washington.

Freeman, R.E. (1984), *Strategic Management: A Stakeholder Approach*, Pitman, Boston, MA, 279 p.

Fryer, K., Antony, J. and Ogden, S. (2009), “Performance management in the public sector”, *International Journal of Public Sector Management*, Vol. 22 No. 6, pp. 478-98.

Grifell-Tatjé, E'., & Lovell, C. A. K. (1995). A note on the Malmquist productivity index. *Economics Letters*, 47,169-175.

Greiling, D. (2006), Performance measurement: a remedy for increasing the efficiency of public services?, *International Journal of Productivity and Performance Management*, Vol. 55 No. 6, pp. 448-65.

Guthrie, J., and Tooley, S. (2007). Reporting performance by New Zealand secondary schools: An analysis of disclosures. *Financial Accountability & Management, 23*(4), 351- 374.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, l4 (3), 351-388.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24(3), 141-1177.

Hanushek, E. A. (2005). The economics of school quality. *German Economic Review, 6*(3), 269-286.

Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Public Analysis and Management* 24(2), 297 -327.

Harrison, J. (2008). Results or process? : alternative approaches to the performance measurement of New Zealand secondary schools. *The University of Auckland PHD Thesis*. Auckland, New Zealand: The University of Auckland.

Harrison, J., and Rouse, P. (2014). Competition and public high school performance. *Socio-Economic Planning Sciences, 48*(1), 10-19.

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

Harrison, J. A., Rouse, P., and De Villiers, C., J. (2012). Accountability and performance measurement: A stakeholder perspective. *Journal of Centrum Cathedra*, *5*(2), 243-258.

Hattie, J. (2002). Schools Like Mine: Cluster Analysis of New Zealand Schools. Auckland, New Zealand: University of Auckland.

Haveman, R., & Wolfe, R. 0995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic literature*, 33(4), 1829-1878.

Hoff, A. (2006). Second stage DEA: Comparison of approaches for modelling the DEA score. *European Journal of Operational Research*, 187, 425-435.

Hood, C. (2000). *Contemporary Public Management: A New Global Paradigm?*. Oxford University Press. doi:10.1093/0198297653.003.0009

Ittner, C. D. and D. F. Larcker (1998), ‘Innovations in Performance Measurement: Trends and Research Implications’, *Journal of Management Accounting Research*, Vol. 10, pp. 205–38.

Jaforullah, M. (2010). Productivity Change in New Zealand Secondary Schools. *Journal of New Business Ideas & Trends, 8(2)*, 14-26.

Jackson, P. M. (1982). The political economy of bureaucracy. London: Philip Allan.

Jensen, B., & Seltzer, A. (2000). Neighbourhood and family effects in educational progress. *The Australian Economic Review, 33*(1), 17-31.

Jones, T.M. (1995), “Instrumental stakeholder theory: a synthesis of ethics and economics”, *Academy of Management Review*, Vol. 20, pp. 404-437.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard- measures that drive performance. *Harvard Business Review, 70(1), 71-79.*

Kaplan, R. S., & Norton, D. P. (1996a). *The balanced scorecard: Translating strategy into action*. Boston, MA: Harvard Business School Press.

Kenny, G. (2003). Strategy: Balanced scorecard: Why it isn’t working. *New Zealand Management, (March).* 32-35.

Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities. In T. C. Koopmans (Ed.). *Activity| Analysis of Production and Allocation, Cowles Commission for Research in Economics, Monograph No. 13*. New York: Wiley.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114, 497-532.

Ladd, H. F. (Ed.). (1996). *Holding schools accountable: Performance-based reform in education.*.Washington, D C: Brookings Institution.

Lapsley, I. (2008). The NPM agenda: back to the future. *Financial Accountability & Management*, Vol. 24 No. 1, pp. 77-96.

Lingle, J. H., & Schiemann, W. A. (1996). From balanced scorecard to strategy gauge: Is measurement worth it. *Management Review (March)*, 56-62.

Lockheed, M. and Hanushek, E. (1994). Concepts of educational efficiency and effectiveness. Human Resources Development and Operations Policy Working Papers, 19 pp.

Malik, A. A. (2011). Performance of New Zealand's secondary schools: A stakeholders' perspective. *PHD Thesis at University of Waikato*, 538.

Mitchell, R. K., Agel, B. R. and Wood, D. J. (1997), "Toward a theory of stakehoder identification and salience", *Academy of Management Review*, Vol. 22, No. 4, pp. 853-866.

Ministry of Education. (1989). *Governing schools: A practical handbook for school trustees,* Wellington, New Zealand: Ministry of Education.

Ministry of Education. (2011). *Statement of intent.* Retrieved from Ministry of Education: http://www.education.govt.nz/assets/Documents/Ministry/Publications/Statements-of-intent/SOI2011Web.pdf

Ministry of Education. (2013). *Enrolment schemes (zoning).* Retrieved from Ministry of Education: http://www.minedu.govt.nz/Parents/AllAges/UsefulInformation/EnrolmentSchemesZones.aspx

Ministry of Education. (2013a). *Teachers' pay, allowances and qualifications.* Retrieved from Ministry of Education.

Ministry of Education. (2014a). *The National Education Guidelines (NEGs).* Retrieved from Ministry of Education: http://www.minedu.govt.nz/theMinistry/EducationInNewZealand/EducationLegislation/TheNationalEducationGuidelinesNEGs.aspx

Ministry of Education. (2014c). *Our strategic direction.* Retrieved from Ministry of Education: http://www.minedu.govt.nz/theMinistry/PublicationsAndResources/StatementOfIntent/SOI2014/OurStrategicDirection.aspx

Ministry of Education. (2014d). *What we do.* Retrieved from Ministry of Education: http://www.minedu.govt.nz/theMinistry/AboutUs/AboutTheMinistry.aspx#sector

Ministry of Education. (2014e). *Performance management.* Retrieved from Ministry of Education: http://www.minedu.govt.nz/Boards/ManagingResources/PerformanceManagement.aspx

Ministry of Education. (2014f). *Performance management.* Retrieved from Ministry of Education: http://www.minedu.govt.nz/NZEducation/EducationPolicies/SchoolEmployment/Employers/ManagingStaff/PerformanceManagement.aspx

Ministry of Education. (2014g). *Chapter 5: Keeping track of performance and effectiveness.* Retrieved 2014, from http://www.minedu.govt.nz/theMinistry/PublicationsAndResources/StatementOfIntent/SOI2007/Chapter_5__Keeping_track_of_performance_and_effectiveness.aspx

Ministry of Education. (2014h). *Interventions: Guide for schools.* Retrieved 2014h, from http://www.minedu.govt.nz/Boards/SupportForBoards/InterventionsInSchools/InterventionsInSchools.aspx

Ministry of Education. (2014i). *The funding, staffing and allowances handbook.* Retrieved 2014, from http://www.minedu.govt.nz/NZEducation/EducationPolicies/Schools/SchoolOperations/Resourcing/ResourcingHandbook.aspx

Ministry of Education. (2014j). Chapter 4: Payments to individuals. Retrieved from Ministry of Education: http://www.minedu.govt.nz/NZEducation/EducationPolicies/Schools/SchoolOperations/Resourcing/ResourcingHandbook/Chapter4.aspx

Ministry of Education. (2014k). Appendix 1: Operational funding rates for 2015 and 2014. Retrieved from Ministry of Education: http://www.minedu.govt.nz/NZEducation/EducationPolicies/Schools/SchoolOperations/Resourcing/ResourcingHandbook/Chapter1/Appendices/Appendix1OperationalFundingRates.aspx

Ministry of Education. (2014l). *New Zealand Education System Overview.* Retrieved from Ministry of Education: http://www.minedu.govt.nz/NZEducation/EducationPolicies/InternationalEducation/ForInternationalStudentsAndParents/NZEdOverview/Key_Education_Agencies_and_Links.aspx

Ministry of Education. (2014m). *How the decile is calculated.* Retrieved from Ministry of Education: http://www.minedu.govt.nz/NZEducation/EducationPolicies/Schools/SchoolOperations/Resourcing/OperationalFunding/Deciles/HowTheDecileIsCalculated.aspx

Ministry of Education. (2015, May 19). *Ministry of Education.* Retrieved from The National Administration Guidelines (NAGs): http://www.education.govt.nz/ministry-of-education/legislation/nags/

Ministry of Education. (2016). *Accounting for donations, fundraising and other income.* Retrieved from Education: http://www.education.govt.nz/school/running-a-school/school-finances/accounting-for-donations-fundraising-and-other-income/#Useofland

Ministry of Education. (2016a). *School deciles.* Retrieved from Ministry of Education: http://www.education.govt.nz/school/running-a-school/resourcing/operational-funding/school-decile-ratings/

Näsi, J. (1995), *Understanding Stakeholder Thinking*, LSR Publications, Helsinki

New Zealand Department of Education. (1988). Tomorrow’s schools: *The reform of education administration in New Zealand*. Wellington, NZ: Government Printer.

New Zealand Qualifications Authority. (2012). *Annual Report on NCEA and New Zealand Scholarship Data and Statistics (2011).*

New Zealand Qualifications Authority. (2013). *Secondary Statistics Consolidated Data Files.*

New Zealand Qualifications Authority. (2014). An Independent Review of the Effectiveness of NZQA’s Implementation of the 2007 NCEA Enhancements. Wellington: New Zealand Qualifications Authority.

New Zealand Qualifications Authority. (2015). History of NCEA. Retrieved from NZQA: http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/history-of-ncea/

New Zealand Qualifications Authority. (2015a). *Literacy and numeracy requirements prior to 2013.* Retrieved from NZQA: http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/subjects/literacy-and-numeracy/level-1-requirements/

New Zealand Qualifications Authority. (2016). NCEA levels and certificates. Retrieved from NZQA: http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/how-ncea-works/ncea-levels-and-certificates/

New Zealand Qualification Authority. (2016a). Factsheet # 9: Are there gender differences in achievement rates? Retrieved from NZQA: http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/the-facts/factsheet-10/

New Zealand Qualification Authority. (2016b). Factsheet # 10: How does NZQA make sure internal assessment is fair and consistent across the country?Retrieved from NZQA: http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/the-facts/factsheet-10/

New Zealand Qualification Authority. (2016c). Factsheet # 12: How are NCEA results recognised overseas?Retrieved from NZQA: http://www.nzqa.govt.nz/qualifications-standards/qualifications/ncea/understanding-ncea/the-facts/factsheet-12/

NZ Herald. (2003, August 11). *Mallard announces $78 million more for education.* Retrieved from nzherald: http://www.nzherald.co.nz/education/news/article.cfm?c_id=35&objectid=3517449

NZ Herald. (2007, March 22). *Your Views: More of your thoughts on NCEA.* Retrieved from nzherald: http://www.nzherald.co.nz/opinion/news/article.cfm?c_id=466&objectid=10429647

NZ Herald. (2009, October 10). Community funding for schools. Retrieved 2014, from The New Zealand Herald: http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10602275

OECD. (2013, November 14). *Government at a Glance.* Retrieved 2014, from OECDiLibrary: http://www.oecd-ilibrary.org/governance/government-at-a-glance-2011_gov_glance-2011-en;jsessionid=6uvlrcor2ialt.x-oecd-live-01

Parliamentary Counsel Office (2010). *Education Act 1989.* Retrieved 2014, from http://www.legislation.govt.nz/act/public/1989/0080/latest/DLM175959.html

Ploom, K. and Haldma, T. (2013). "Balanced performance management in the public education system", *Baltic Journal of Management*, Vol. 8 No.2, pp.183-207.

Pollitt, C. and Summa, H. (1997). Trajectories of reform: management change in four countries. *Public Money and Management*, Vol. 17 No. 1, pp. 7-18.

Ramanathan, K. V. (1985). A proposed framework for designing management control systems in not-for-profit organizations. *Financial Accountability & Management in Government, Public Services, and Charities*, 1 (1), 75-92.

Rigby, D., & Bilodeau, B. (2009). *Management tools and trends 2009.* Retrieved 2014, from http://www.bain.com/bainweb/PDFs/cms/Public/Management_Tools_2009.pdf

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools and academic achievement. *Econometrica*, 73(2), 417 -458.

Rouse, P., & Putterill, M. (2003). An integral framework for performance measurement. *Management Decision*, *41*(8), 791-805.

Ruggiero, J. (1996). On the measurement of technical efficiency in the public sector. *European Journal of Operational Research*, 90(3), 553-565.

Ruggiero, J. (1999). Nonparametric analysis of educational costs. *European Journal of Operational Research*, 119(3), 605-612.

Ruggiero, J. (2003). Comment on estimating school efficiency. *Economics of Education review*, 22(6), 631-634.

Ruggiero, J. (2006), Measurement error, education production and data envelopment analysis. *Economics of Education Review*, 25,327-333.

Scott, D. (2008), *How Does Achievement at School Effect Achievement in Tertiary Education?*, Tertiary Sector Performance Analysis and Reporting, Ministry of Education, Wellington.

Shulruf, B., Hattie, J., & Tumen, S. (2010). New Zealand's standard-based assessment for secondary schools (NCEA): Implications for policy makers. *Asia Pacific Journal of Education*, 30(2), 141-165.

Smelt, S. (1998). *Today"s schools: Governance and quality*, Wellington, New Zealand: Institute of Policy Studies.

Statistics New Zealand. (2015). Implicit Price Deflators: Annual: Gross National Expenditure. Retrieved 30 November, 2015, from New Zealand Times Series Database, Reference SNZ51.902

Taut, S. (2008). What have we learned about stakeholder involvement in program evaluation? *Studies in Educational Evaluation*, 34(4), 224-230.

Thomas, S., Peng, W. J., & Gray, J. (2007). Modelling Patterns of Improvement over Time: Value Added Trends in English Secondary School Performance across Ten Cohorts. *Oxford Review of Education*, (3). 261.

Trading Economics. (2014). Gross national expenditure deflator (base year varies by country) in New Zealand. Retrieved from Trading Economics: http://www.tradingeconomics.com/new-zealand/gross-national-expenditure-deflator-base-year-varies-by-country-wb-data.html

Wöbmann, L. et al. (2007), School Accountability, Autonomy, Choice, and the Level of Student Achievement: International Evidence from PISA 2003. OECD Education Working Papers, No. 13, OECD Publishing.

Zhang, Y., & Bartels, R. (1998). The Effect of Sample Size on the Mean Efficiency in DEA with an Application to Electricity Distribution in Australia, Sweden and New Zealand. *Journal of Productivity Analysis*, 9(3), 187-204.

Zhu, J. (2003). *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver*. Kluwer Academic Publishers, Boston.

Appendices

Appendix I Description of control linkages under Ramanathan's framework

This framework incorporates the disciplines related to both management control and performance measurement and decomposes benefit-cost criteria into a number of interrelated control links:

$$\frac{B}{C} = \left(\frac{B}{OC}\right)\left(\frac{OC}{O}\right)\left(\frac{O}{I}\right)\left(\frac{I}{C}\right)$$

Where,

- Social benefit (B): a financial measure of the social benefit generated by the organization to satisfy the social needs.
- Outcomes (OC): non-financial measures of social benefit. Multiple outcome indicators should reflect multiple social objectives and are connected with social benefits so as to ensure public accountability.
- Outputs (O): non-financial measures of organization's activity which should be linked to social outcomes.
- Inputs (I): non-financial measures of the resources spent by the organization, which should be linked to the production of outputs.
- Costs (C): financial measures used up by the organization
- Social benefit per dollar spent (B/C)
- Social benefit per successful outcome (B/OC)
- Success rate in organization (OC/O)
- Productivity rate (O/I)
- Resources available per dollar spent (I/C)

Source: Ramanathan (1985)

Appendix II Summary of expenditures per intermediate student

Table A2.1 Summary of expenditures per intermediate student[95]

Year	Teacher $/student		Other $/student	
	Roll<500	Roll >=500	Roll<500	Roll>=500
2004	3680	3042	1263	947
2005	3902	3208	1348	1000
2006	4158	3385	1430	1046
2009	4938	3914	3216	2379
2010	5062	4107	3228	2361
2011	5242	4232	3391	2495

[95] All the expenditures per student in this table are not indexed to adjust inflation effect because each of them is used for expenditure splitting in specific year. After the splitting, the adjusted expenditures would be indexed to form pooled data.

Appendix III Correlations between roll and school passes by decile

Table A3.1 Descriptive statistics for NCEA L1-3 pass rate participation by decile 2009-2011

					95% Confidence Interval for Mean	
Decile	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound
1	1580	.337128	.3875504	.0097499	.318004	.356253
2	1652	.368177	.3598860	.0088544	.350810	.385544
3	1540	.386706	.3721664	.0094837	.368104	.405308
4	1660	.403758	.3599376	.0088343	.386430	.421085
5	1540	.435544	.3690393	.0094040	.417098	.453990
6	1792	.442403	.3754179	.0088684	.425009	.459796
7	1428	.447047	.3878769	.0102643	.426912	.467182
8	1584	.452903	.3880335	.0097497	.433780	.472027
9	1308	.482132	.3983817	.0110153	.460523	.503742
10	1592	.490179	.4317518	.0108209	.468954	.511404
Total	15676	.423383	.3857370	.0030809	.417344	.429422

Table A3.2 Descriptive statistics for NCEA L1-3 pass rate participation by decile 2004-2006

					95% Confidence Interval for Mean	
Decile	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound
1	1380	.278773	.3510016	.0094486	.260238	.297308
2	1520	.304784	.3177439	.0081500	.288797	.320770
3	1260	.333381	.3296212	.0092860	.315163	.351599
4	1612	.358474	.3298523	.0082156	.342360	.374588
5	1692	.348758	.3320719	.0080729	.332924	.364592
6	1676	.384649	.3497563	.0085434	.367892	.401405
7	1560	.398546	.3590803	.0090914	.380713	.416378
8	1268	.411773	.3607475	.0101308	.391898	.431648
9	1112	.459114	.3678480	.0110310	.437470	.480758
10	1416	.484501	.4021185	.0106862	.463538	.505463
Total	14496	.373973	.3546724	.0029458	.368199	.379747

Appendix IV Skewness and kurtosis statistics for DEA variables

Table A4.1 Skewness and kurtosis statistics - DEA model variables for 2009-2011

	Full sample (n=865)				Pure secondary school sample (n=615)			
	Skewness		Kurtosis		Skewness		Kurtosis	
	Stat	Std. Err.	Stat	Std. Err.	Stat	Std. Err.	Stat	Std. Err.
SES decile	-0.06	0.083	-1.004	0.166	-0.009	0.099	-0.978	0.197
Indexed teachers$	1.213	0.083	1.401	0.166	0.992	0.099	0.861	0.197
Indexed other$	1.438	0.083	2.912	0.166	1.381	0.099	2.739	0.197
Roll (years 9-15)	1.222	0.083	1.276	0.166	0.994	0.099	0.642	0.197
# NCEA Level 1 passes	1.314	0.083	1.990	0.166	1.097	0.099	1.339	0.197
# NCEA Level 2 passes	1.450	0.083	2.631	0.166	1.228	0.099	1.831	0.197
# NCEA Level 3 passes	1.872	0.083	5.134	0.166	1.664	0.099	3.879	0.197
# UE	1.967	0.083	5.596	0.166	1.745	0.099	4.185	0.197
#NCEA L3 M&E	2.952	0.083	12.640	0.166	2.745	0.099	10.307	0.197

Table A4.2 Skewness and kurtosis statistics - DEA model variables for 2004-2006

	Full sample (n=856)				Pure secondary school sample (n=611)			
	Skewness		Kurtosis		Skewness		Kurtosis	
	Stat	Std. Err.	Stat	Std. Err.	Stat	Std. Err.	Stat	Std. Err.
SES decile	-0.019	0.084	-1.017	0.167	0.035	0.099	-0.988	0.197
Indexed teachers$	1.301	0.084	2.293	0.167	1.144	0.099	2.026	0.197
Indexed other$	1.370	0.084	2.078	0.167	1.256	0.099	1.647	0.197
Roll (years 9-15)	1.301	0.084	1.846	0.167	1.127	0.099	1.365	0.197
# NCEA Level 1 passes	1.717	0.084	4.264	0.167	1.534	0.099	3.442	0.197
# NCEA Level 2 passes	1.992	0.084	5.541	0.167	1.768	0.099	4.254	0.197
# NCEA Level 3 passes	2.374	0.084	7.792	0.167	2.114	0.099	5.858	0.197
# UE	2.420	0.084	8.344	0.167	2.172	0.099	6.403	0.197
# NCEA L3 M&E	3.255	0.084	14.228	0.167	2.938	0.099	11.054	0.197

Appendix V Comparison of other expenditures between two periods

Table A5.1 Comparison of other operating expenditures between 2004-2005 and 2010-2011

	2004-2005		2010-2011		*t*-test		Kruskal-Wallis U-test	
	Mean	**St. Dev**	**Mean**	**St. Dev**	***t*-value**	**Effect size**	***z*-value**	**Effect size**
InxAdmin$ 2198	372,740	216,703	409,754	245,544	-2.84**	0.080	2.775**	0.078
InxProperty$ 2595	379,045	213,147	1,454,795	781,139	-33.50***	0.778	28.402***	0.801
InxLearning$ excl teacher$ 2398	469,089	381,571	522,050	376,166	-2.48*	0.070	3.537***	0.100

Table A5.2 Comparison of property expenses between 2004-2005 and 2010-2011

	2004-2005		2010-2011		*t*-test		Kruskal-Wallis U-test	
	Mean	**St. Dev**	**Mean**	**St. Dev**	***t*-value**	**Effect size**	***z*-value**	**Effect size**
InxContract service 2520	13,508	36,229	32,990	58,491	-7.17***	0.21	8.01	0.22
Inxenergy$ 2550	61,422	34,015	82,082	44,853	-9.29***	0.26	9.43	0.26
InxPersonnel 2570	106,326	74,140	118,230	87,117	-2.63**	0.07	2.325	0.07
InxUseofland& Building 2585	0	0	1,190,248	701,002	-43.06***	0.86	32.64	0.91

Appendix VI Full sample: cost per student and efficiency by school size

Table A6.1 Full school sample: cost per student and efficiency by school size for 2009-2011

Roll Size[96]	Count	Indexed cost per student		Categorical VRS cost efficiency		Categorical VRS academic efficiency		Categorical VRS academic effectiveness		Categorical VRS overall efficiency	
		Mean	St Dev	Mean	St Dev	Mean	St Dev	Mean	St Dev	Mean	St Dev
<=300	86	11647.89	1269.42	0.72	0.14	0.86	0.13	0.81	0.18	0.60	0.27
301-600	261	9341.37	1305.16	0.72	0.11	0.81	0.11	0.73	0.17	0.53	0.23
601-900	219	8000.39	968.63	0.77	0.09	0.83	0.11	0.74	0.15	0.57	0.21
901-1200	104	7421.25	648.48	0.81	0.07	0.86	0.08	0.71	0.13	0.60	0.17
1,201-1,500	87	6847.91	500.95	0.87	0.07	0.87	0.08	0.74	0.11	0.64	0.17
1,501-1,800	50	6779.82	433.43	0.86	0.06	0.88	0.10	0.80	0.11	0.70	0.15
1,800-2,100	26	6485.12	303.35	0.92	0.05	0.91	0.12	0.85	0.10	0.77	0.17
2,101-2,400	15	6537.54	512.32	0.92	0.05	0.88	0.21	0.87	0.13	0.78	0.22
2,401+	17	6812.37	549.87	0.95	0.04	0.74	0.32	0.87	0.16	0.69	0.34
Total	865	8416.97	1759.27	0.78	0.12	0.84	0.12	0.75	0.16	0.59	0.22

[96] Roll size here is the number of secondary students (years 9-15) in each school.

Table A6.2 Full school sample: cost per student and efficiency by school size for 2004-2006

Roll Size	Count	Indexed cost per student		Categorical VRS cost efficiency		Categorical VRS academic efficiency		Categorical VRS academic effectiveness		Categorical VRS overall efficiency	
		Mean	St Dev	Mean	St Dev	Mean	St Dev	Mean	St Dev	Mean	St Dev
<=300	99	8670.81	966.52	0.74	0.13	0.77	0.18	0.79	0.17	0.54	0.30
301-600	259	6999.23	946.44	0.78	0.11	0.73	0.15	0.79	0.13	0.50	0.24
601-900	193	6098.41	676.79	0.83	0.09	0.74	0.15	0.82	0.13	0.55	0.23
901-1200	128	5834.96	617.22	0.84	0.08	0.78	0.12	0.83	0.11	0.59	0.19
1,201-1,500	93	5458.75	420.64	0.90	0.07	0.81	0.11	0.83	0.10	0.62	0.20
1,501-1,800	33	5330.46	320.73	0.90	0.06	0.79	0.11	0.84	0.11	0.65	0.18
1,800-2,100	19	5197.29	287.77	0.95	0.05	0.87	0.13	0.91	0.07	0.78	0.17
2,101-2,400	18	5276.62	366.05	0.95	0.04	0.89	0.12	0.90	0.08	0.73	0.16
2,401+	14	5213.10	277.03	0.98	0.03	0.83	0.24	0.96	0.07	0.79	0.28
Total	856	6478.22	1247.25	0.82	0.11	0.76	0.15	0.82	0.13	0.56	0.24

Appendix VII Full sample: VRS efficiency scores by school decile

Table A7.1 Full sample: VRS efficiency scores by school decile for 2009-2011

SES	N	Standard model VRS			Categorical VRS			Standard model VRS			Categorical VRS			Standard model VRS			Categorical VRS			Standard model VRS			Categorical VRS		
		# Efficient	Cost efficiency		# Efficient	Cost efficiency		# Efficient	Academic efficiency		# Efficient	Academic efficiency		# Efficient	Academic effectiveness		# Efficient	Academic effectiveness		# Efficient	Overall efficiency		# Efficient	Overall efficiency	
			Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.
1	39	0	.67	.12	6	.87	.12	0	.66	.11	12	.86	.13	0	.38	.10	10	.82	.15	0	.16	.10	9	.58	.31
2	81	1	.65	.11	7	.77	.12	1	.67	.10	7	.80	.12	1	.49	.14	18	.78	.18	1	.19	.12	13	.55	.28
3	78	0	.67	.10	6	.78	.12	1	.70	.11	10	.79	.13	2	.55	.14	23	.83	.17	0	.24	.12	9	.55	.27
4	98	2	.69	.12	5	.77	.11	1	.72	.10	8	.81	.11	2	.58	.15	8	.80	.13	0	.29	.13	6	.54	.21
5	107	3	.75	.13	8	.79	.13	3	.77	.11	9	.83	.12	4	.63	.14	6	.66	.16	2	.38	.17	8	.58	.22
6	126	4	.76	.10	3	.76	.10	4	.78	.10	8	.82	.10	1	.63	.11	7	.68	.14	0	.41	.14	5	.57	.17
7	89	2	.78	.10	2	.78	.10	3	.84	.08	9	.89	.08	4	.71	.12	5	.73	.13	2	.54	.17	5	.64	.20
8	99	1	.79	.11	1	.79	.11	2	.83	.10	8	.86	.10	2	.72	.12	3	.73	.12	3	.54	.16	4	.60	.18
9	92	1	.78	.12	1	.78	.12	3	.87	.08	4	.89	.08	8	.78	.12	8	.78	.12	6	.64	.18	6	.66	.18
10	56	1	.75	.13	1	.75	.13	8	.85	.19	8	.85	.19	6	.84	.11	6	.84	.11	5	.68	.22	5	.68	.22
Total	865	15	.74	.12	40	.78	.12	26	.77	.13	83	.84	.12	30	.64	.17	94	.75	.16	19	.42	.22	70	.59	.22

Table A7.2 Full sample: VRS efficiency scores by school decile for 2004-2006

SES	N	Standard model VRS			Categorical VRS			Standard model VRS			Categorical VRS			Standard model VRS			Categorical VRS			Standard model VRS			Categorical VRS		
		# Efficient	Cost efficiency		# Efficient	Cost efficiency		# Efficient	Academic efficiency		# Efficient	Academic efficiency		# Efficient	Academic effectiveness		# Efficient	Academic effectiveness		# Efficient	Overall efficiency		# Efficient	Overall efficiency	
			Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.		Mean	S.D.
1	34	0	.68	.10	9	.89	.10	0	.48	.17	9	.75	.22	2	.60	.19	13	.82	.19	0	.10	.07	8	.57	.33
2	82	2	.73	.11	12	.83	.11	3	.55	.16	8	.73	.18	3	.65	.15	14	.76	.17	3	.22	.21	8	.46	.26
3	74	0	.76	.11	9	.82	.11	0	.56	.13	10	.73	.19	4	.72	.13	14	.83	.15	0	.22	.11	10	.52	.28
4	103	0	.74	.11	4	.80	.11	1	.60	.12	4	.70	.14	5	.72	.14	7	.79	.14	0	.26	.14	4	.50	.24
5	107	4	.76	.12	7	.79	.11	0	.64	.11	4	.72	.13	1	.76	.12	8	.83	.12	0	.33	.14	6	.51	.21
6	107	1	.81	.11	4	.83	.10	2	.68	.11	3	.75	.12	0	.73	.12	4	.80	.13	1	.38	.16	4	.55	.22
7	103	5	.82	.10	5	.83	.09	0	.68	.12	4	.77	.14	2	.78	.10	2	.80	.10	3	.46	.19	6	.58	.22
8	89	2	.83	.10	3	.83	.11	1	.74	.09	4	.82	.09	3	.80	.09	3	.81	.09	2	.48	.16	3	.59	.19
9	73	2	.83	.10	2	.83	.10	3	.79	.10	10	.86	.10	2	.85	.08	7	.86	.08	2	.60	.20	7	.69	.22
10	84	4	.82	.14	3	.82	.14	12	.83	.13	12	.83	.13	13	.89	.08	12	.89	.08	8	.70	.19	8	.70	.19
Total	856	20	.79	.12	58	.82	.11	22	.67	.15	68	.76	.15	35	.76	.14	84	.82	.13	19	.39	.23	64	.56	.24

Appendix VIII Statistics of efficient schools (VRS categorical)

Table A8.1 Full sample: statistics of efficient schools for categorical VRS model for 2009-2011

Model	# Efficient schools	SES Decile										School size	Type		Ownership		Gender		
		1	2	3	4	5	6	7	8	9	10	Mean (Std. Dev.)	Expanded schools (n=237)	Pure secondary schools (n=628)	Integrated schools (n=174)	Non-integrated schools (n=691)	Boy schools (n=117)	Girl schools (n=140)	Co-educational schools (n=608)
Cost efficient	40	6	7	6	5	8	3	2	1	1	1	964.53 (839.47)	18 (45%)	22 (55%)	14 (35%)	26 (65%)	4 (10%)	6 (15%)	30 (75%)
Academic efficient	83	12	7	10	8	9	8	9	8	4	8	1067.37 (843.23)	26 (31%)	57 (69%)	26 (31%)	57 (69%)	4 (5%)	19 (23%)	60 (72%)
Academic effective	94	10	18	23	8	6	7	5	3	8	6	840.71 (807.70)	31 (33%)	63 (67%)	24 (26%)	70 (74%)	3 (3%)	24 (26%)	67 (71%)
CEc&AEc	10	3	1	1	1	2	0	0	1	0	1	1068.60 (1119.26)	4 (40%)	6 (60%)	2 (20%)	8 (80%)	1 (10%)	1 (10%)	8 (80%)
CEc&AEt	16	3	1	4	4	2	0	0	1	0	1	1164.38 (1113.32)	6 (38%)	10 (62%)	3 (19%)	13 (81%)	1 (6%)	2 (13%)	13 (81%)
AEc&AEt	37	6	4	8	4	2	2	5	2	1	3	1008.59 (961.53)	17 (46%)	20 (54%)	14 (38%)	23 (62%)	1 (2%)	8 (22%)	28 (76%)
CEc, AEc&AEt	9	2	1	1	1	2	0	0	1	0	1	1154.11 (1151.99)	4 (44%)	5 (56%)	2 (22%)	7 (78%)	1 (11%)	1 (11%)	7 (78%)
Overall	70	9	13	9	6	8	5	5	4	6	5	881.11 807.72	30 (43%)	40 (57%)	26 (37%)	44 (63%)	3 (4%)	23 (33%)	44 (63%)

Table A8.2 Pure secondary school sample: statistics of efficient schools for categorical VRS model for 2009-2011

Model	# Efficient schools	SES Decile										School size	Ownership		Gender		
		1	2	3	4	5	6	7	8	9	10	Mean Std. Dev.	Integrated schools (n=60)	Non-integrated schools (n=555)	Boy schools (n=80)	Girl schools (n=102)	Co-educational schools (n=433)
Cost efficient	31	6	8	4	3	2	3	0	3	1	1	1228.97 (849.84)	8 (26%)	23 (74%)	1 (3%)	7 (23%)	23 (74%)
Academic efficient	68	11	7	8	5	7	6	5	7	6	6	1245.57 (829.21)	17 (25%)	51 (75%)	2 (3%)	19 (28%)	47 (69%)
Academic effective	78	11	15	20	6	4	4	5	3	6	4	1003.32 (827.884)	12 (15%)	66 (85%)	2 (3%)	22 (28%)	54 (69%)
CEc&AEc	12	4	4	1	0	0	0	0	1	1	1	970.33 (1027.66)	6 (50%)	6 (50%)	0	3 (25%)	9 (75%)
CEc&AEt	16	3	4	3	3	0	0	0	1	1	1	1263.75 (1043.55)	6 (38%)	10 (62%)	1 (6%)	5 (31%)	10 (63%)
AEc&AEt	28	5	4	6	2	1	1	2	2	3	2	1280.82 (966.911)	9 (32%)	19 (68%)	0	9 (32%)	19 (68%)
CEc, AEc&AEt	10	3	3	1	0	0	0	0	1	1	1	1100.70 (1084.973)	5 (50%)	5 (50%)	0	3 (30%)	7 (70%)
Overall	49	5	8	8	2	4	3	7	3	5	4	1181.88 (852.37)	13 (27%)	36 (73%)	0	23 (47%)	26 (53%)

Table A8.3 Full sample: statistics of efficient schools under categorical VRS model for 2004-2006

Model	# Efficient schools	SES Decile										School size	Type		Ownership		Gender		
		1	2	3	4	5	6	7	8	9	10	Mean (Std. Dev.)	Expended schools (n=229)	Pure secondary schools (n=627)	Integrated schools (n=183)	Non-integrated schools (n=673)	Boy schools (n=122)	Girl schools (n=146)	Co-educational schools (n=588)
Cost efficient	58	9	12	9	4	7	4	5	3	2	3	1132.67 (845.73)	17 (29%)	41(71%)	16 (28%)	42(72%)	6(11%)	6 (11%)	46 (78%)
Academic efficient	68	9	8	10	4	4	3	4	4	10	12	998.88 (853.41)	19(28%)	49(72%)	27(40%)	41(60%)	9(12%)	25(38%)	34(50%)
Academic effective	84	13	14	14	7	8	4	2	3	7	12	944.32 (804.86)	27(32%)	57(68%)	21(25%)	63(75%)	11(13%)	15(18%)	58(69%)
CEc&AEc	14	2	3	2	2	1	1	0	0	1	2	1478.57 (1210.82)	2(14%)	12(86%)	4(29%)	10(71%)	1(7%)	2(14%)	11(79%)
CEc&AEt	17	3	3	4	3	1	0	0	0	1	2	1280.06 (1121.39)	5(29%)	12(71%)	3(18%)	14(82%)	0(0%)	2(12%)	15(88%)
AEc&AEt	28	6	4	5	3	0	0	1	0	4	5	1357.86 (1011.01)	5(18%)	23(82%)	6(21%)	22(79%)	3(11%)	6(21%)	19(68%)
CEc, AEc&AEt	9	2	1	1	2	0	0	0	0	1	2	1796.44 (1219.64)	1(11%)	8(89%)	1(11%)	8(89%)	0(%)	1(11%)	8(89%)
Overall	64	8	8	10	4	6	4	6	3	7	8	951.67 (826.63)	24(38%)	40(62%)	28(44%)	36(56%)	5(8%)	22(34%)	37(58%)

Table A8.4 Pure secondary school sample: statistics of efficient schools under categorical VRS model for 2004-2006

Model	# Efficient schools	SES Decile										School size	Ownership		Gender		
		1	2	3	4	5	6	7	8	9	10	Mean (Std. Dev.)	Integrated schools (n=64)	Non-integrated schools (n=547)	Boy schools (n=84)	Girl schools (n=100)	Co-educational schools (n=427)
Cost efficient	42	6	9	5	2	6	2	5	2	2	3	1425.64 815.34	6 (14%)	36(86%)	3(7%)	5 (12%)	34 (81%)
Academic efficient	55	6	5	10	3	3	3	8	4	7	6	1173.05 862.55	18(33%)	37(67%)	5(9%)	20(36%)	30(55%)
Academic effective	82	12	14	12	9	7	3	6	2	9	8	1005.88 785.67	12(15%)	70(85%)	7(8%)	18(22%)	57(70%)
CEc&AEc	11	1	1	2	2	1	0	1	0	1	2	1840.55 1111.51	2(18%)	9(82%)	1(9%)	1(9%)	9(82%)
CEc&AEt	15	5	1	3	2	0	0	1	0	1	2	1527.73 1044.50	3(20%)	12(80%)	0(%)	3(20%)	12(80%)
AEc&AEt	28	5	3	5	3	1	0	2	0	5	4	1355.04 1012.77	7(25%)	21(75%)	2(7%)	8(29%)	18(64%)
CEc, AEc&AEt	8	1	0	1	2	0	0	1	0	1	2	2008.25 1114.80	1(13%)	7(87%)	0(%)	1(12%)	7(88%)
Overall	50	7	6	7	4	4	1	4	3	9	5	1132.92 854.76	17(34%)	33(66%)	3(6%)	23(46%)	24(48%)

Appendix IX Comparison of efficient schools to total sample

Table A9.1 Full sample: comparison of cost efficient schools to total sample

Test period: 2009-2011						
SES	Full sample			Cost efficient schools		
	N	Roll Mean	Inx operational expenditure/student Mean	N	Roll mean	Inx operational expenditure/student Mean
1	39	711.26	9054.08	6	651.50	8722.92
2	81	605.05	9512.25	7	1010.43	8074.49
3	78	573.95	9395.73	6	578.67	8914.15
4	98	728.17	9150.40	5	1336.00	8685.35
5	107	806.07	8377.47	8	726.63	8372.51
6	126	828.80	7987.73	3	1183.33	5975.59
7	89	877.34	7779.83	2	575.00	6268.28
8	99	1044.57	7651.03	1	2596	6256.44
9	92	1031.13	7780.16	1	1292	5691.19
10	56	1199.00	8196.44	1	3046	6895.35
Total	865	840.53	8416.97	40	964.53	8051.43
Test period: 2004-2006						
SES	Full sample			Cost efficient schools		
	N	Roll Mean	Inx operational expenditure/student Mean	N	Roll Mean	Inx operational expenditure/student Mean
1	34	632.53	7290.83	9	716.44	6841.60
2	82	581.41	7245.24	12	709.33	6718.49
3	74	696.00	6742.44	9	1006.56	5755.50
4	103	716.46	6795.65	4	1355.50	6137.09
5	107	786.54	6684.24	7	1606.57	5342.24
6	107	831.40	6222.24	4	812.75	5347.44
7	103	857.95	6048.33	5	1194.00	4900.88
8	89	815.46	6142.80	3	1053.67	5950.67
9	73	1048.01	5896.28	2	2575.00	4898.81
10	84	1154.26	6230.49	3	2492.00	5196.68
Total	856	820.10	6478.22	58	1132.67	5949.54

Table A9.2 Full sample: comparison of academic efficient schools to total sample

Test period: 2009-2011										
	Full sample					Academic efficient schools				
SES	N	Roll Mean	NCEAL1 pass rate Mean	NCEAL2 pass rate Mean	NCEAL3 pass rate Mean	N	Roll Mean	NCEAL1 pass rate Mean	NCEAL2 pass rate Mean	NCEAL3 pass rate Mean
1	39	711.26	.1586	.1299	.0502	12	820.25	.1768	.1500	.0675
2	81	605.05	.1596	.1262	.0484	7	921.57	.1809	.1561	.0713
3	78	573.95	.1686	.1328	.0546	10	734.40	.2054	.1604	.0830
4	98	728.17	.1736	.1396	.0658	8	1349.13	.1979	.1557	.0737
5	107	806.07	.1812	.1505	.0737	9	963.78	.2087	.1646	.0958
6	126	828.80	.1822	.1563	.0849	8	565.63	.2213	.1993	.1393
7	89	877.34	.1924	.1678	.1027	9	993.67	.2020	.1731	.1378
8	99	1044.57	.1916	.1658	.0987	8	1785.00	.2142	.1868	.1155
9	92	1031.13	.1981	.1777	.1191	4	1195.00	.2306	.2033	.1467
10	56	1199.00	.1891	.1722	.1262	8	1619.88	.2103	.1987	.1607
Total	865	840.53	.1810	.1531	.0834	83	1067.37	.2026	.1720	.1053
Test period: 2004-2006										
	Full sample					Academic efficient schools				
SES	N	Roll Mean	NCEAL1 pass rate Mean	NCEAL2 pass rate Mean	NCEAL3 pass rate Mean	N	Roll Mean	NCEAL1 pass rate Mean	NCEAL2 pass rate Mean	NCEAL3 pass rate Mean
1	34	632.53	0.125	0.0876	0.0201	9	609.22	0.1463	0.104	0.0304
2	82	581.41	0.1447	0.0986	0.0347	8	402.5	0.2016	0.1546	0.0774
3	74	696.00	0.1453	0.1001	0.038	10	1058.3	0.1813	0.1177	0.0615
4	103	716.46	0.1545	0.1157	0.0474	4	1800.5	0.2061	0.1671	0.0843
5	107	786.54	0.1628	0.1235	0.0564	4	994.5	0.2259	0.1722	0.0948
6	107	831.40	0.1759	0.1337	0.063	3	689.67	0.2612	0.1937	0.1058
7	103	857.95	0.1704	0.1373	0.0737	4	967	0.1798	0.1857	0.1098
8	89	815.46	0.1871	0.1532	0.0742	4	973.25	0.2211	0.1939	0.0993
9	73	1048.01	0.1941	0.163	0.094	10	1045.6	0.2061	0.1874	0.1501
10	84	1154.26	0.1959	0.1783	0.1124	12	1431	0.2156	0.2299	0.1312
Total	856	820.10	0.1681	0.1315	0.0637	68	998.88	0.1986	0.1682	0.0949

Table A9.3 Full sample: comparison of academic effective schools to total sample

Test period: 2009-2011								
	Full sample				Academic effective schools			
SES	N	Roll Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean	N	Roll Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean
1	39	711.26	0.0322	0.0023	10	720.5	0.0431	0.0048
2	81	605.05	0.0365	0.0046	18	610.44	0.0445	0.0079
3	78	573.95	0.0455	0.0073	23	630.22	0.0616	0.0119
4	98	728.17	0.0549	0.0114	8	1316.12	0.0752	0.0223
5	107	806.07	0.0649	0.0152	6	537.33	0.072	0.0282
6	126	828.80	0.0725	0.019	7	328	0.1052	0.0364
7	89	877.34	0.0926	0.0293	5	821.2	0.1347	0.05
8	99	1044.57	0.0902	0.0289	3	2257.67	0.121	0.0634
9	92	1031.13	0.1118	0.0407	8	1135.62	0.1572	0.0792
10	56	1199.00	0.1229	0.0468	6	1721	0.1354	0.0715
Total	865	840.53	0.0735	0.0208	94	840.71	0.0801	0.0273
Test period: 2004-2006								
	Full sample				Academic effective schools			
SES	N	Roll Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean	N	Roll Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean
1	34	632.53	0.0171	0.0011	13	655.15	0.0205	0.0021
2	82	581.41	0.0296	0.0029	14	517	0.0438	0.0054
3	74	696	0.0341	0.0058	14	684.43	0.0445	0.0089
4	103	716.46	0.0435	0.0083	7	1368.71	0.0714	0.0269
5	107	786.54	0.0522	0.0114	8	497.25	0.0553	0.0184
6	107	831.40	0.057	0.0119	4	751.75	0.0805	0.0151
7	103	857.95	0.068	0.0171	2	803	0.1015	0.028
8	89	815.46	0.0707	0.0167	3	511	0.0971	0.0246
9	73	1048.01	0.0901	0.0253	7	1641.86	0.1187	0.0407
10	84	1154.26	0.1088	0.0368	12	1899	0.1191	0.0523
Total	856	820.10	0.0593	0.0143	84	944.32	0.0657	0.0198

Table A9.4 Full sample: comparison of overall efficient schools to total sample

Test period: 2009-2011										
	Full sample					Overall efficient schools				
SES	N	Roll Mean	Inx operational expenditure /student Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean	N	Roll Mean	Inx operational expenditure /student Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean
1	39	711.26	9054.08	0.0322	0.0023	9	578.22	9283.86	0.0396	0.0056
2	81	605.05	9512.25	0.0365	0.0046	13	507.77	9843.56	0.0506	0.0091
3	78	573.95	9395.73	0.0455	0.0073	9	766.78	8502.48	0.0717	0.0207
4	98	728.17	9150.40	0.0549	0.0114	6	869	10301.98	0.0702	0.0198
5	107	806.07	8377.47	0.0649	0.0152	8	761.25	8931.13	0.0827	0.0257
6	126	828.80	7987.73	0.0725	0.019	5	545.4	8377.94	0.1237	0.0528
7	89	877.34	7779.83	0.0926	0.0293	5	1208.2	6799.16	0.1213	0.063
8	99	1044.57	7651.03	0.0902	0.0289	4	1488.75	7018.49	0.13	0.0565
9	92	1031.13	7780.16	0.1118	0.0407	6	1410.17	7336.90	0.1503	0.0723
10	56	1199.00	8196.44	0.1229	0.0468	5	1696.8	7351.72	0.1549	0.0683
Total	865	840.53	8416.97	0.0735	0.0208	70	881.11	8657.77	0.0881	0.0323
Test period: 2004-2006										
	Full sample					Overall efficient schools				
SES	N	Roll Mean	Inx operational expenditure /student Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean	N	Roll Mean	Inx operational expenditure /student Mean	UE pass rate Mean	NCEA L3 M&E pass rate Mean
1	34	632.53	7290.83	0.0171	0.0011	8	725.75	7140.29	0.0194	0.0033
2	82	581.41	7245.24	0.0296	0.0029	8	422.88	7993.57	0.0599	0.0057
3	74	696.00	6742.44	0.0341	0.0058	10	962.7	6478.77	0.0586	0.0123
4	103	716.46	6795.65	0.0435	0.0083	4	1399	6146.04	0.0772	0.0262
5	107	786.54	6684.24	0.0522	0.0114	6	1138.67	5613.50	0.0799	0.0255
6	107	831.40	6222.24	0.057	0.0119	4	346	6081.56	0.0891	0.0162
7	103	857.95	6048.33	0.068	0.0171	6	709	5872.69	0.1159	0.0336
8	89	815.46	6142.80	0.0707	0.0167	3	506.67	6050.95	0.1024	0.0227
9	73	1048.01	5896.28	0.0901	0.0253	7	1431.71	6219.53	0.1331	0.041
10	84	1154.26	6230.49	0.1088	0.0368	8	1560.38	5585.25	0.1437	0.0668
Total	856	820.10	6478.22	0.0593	0.0143	64	951.67	6407.15	0.0851	0.0251

Appendix X Non-radial slacks for model variables

Table A10.1 Non-radial slacks for model variables (categorical VRS model) period 2009-2011

	Full sample (n=865)		Pure secondary school sample (n=615)		Cost efficiency model		Academic efficiency model		Academic effectiveness model		Overall efficiency model	
					FS[97] (n=865)	PS[98] (n=615)	FS (n=865)	PS (n=615)	FS (n=865)	PS (n=615)	FS (n=865)	PS (n=615)
	Mean	Std Dev.	Mean	Std Dev.	# Units[99] with non-zero slack[100] (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)
Teachers' salaries	4,054,814	2,099,730	4,685,385	2,129,390	0	0					0	0
Other expenditures	2,351,543	1,259,029	2,703,756	1,268,228	0	0					0	0
Roll	841	540	990	554	0	0	0	0				
# NCEA L1 passes	151	100	177	104			58 (6.7%)	122 (19.9%)	0	0		
# NCEA L2 passes	130	91	152	96			225 (26%)	179 (29.1%)	0	0		
# NCEA L3 passes	73	61	84	66			358 (41.4%)	263 (42.8%)	0	0		
# UE passes	65	58	75	64					7 (0.8%)	4 0.7%	22 2.5%	35 5.7%
# NCEA L3 M&E	20	26	23	29					188 (21.7%)	133 21.6%	657 75.8%	437 71.1%

[97] FS denotes the full sample.
[98] PS represents the pure secondary school sample.
[99] A unit refers to a school year.
[100] Zero slack is at 0.01 level.

Table A10.2 Non-radial slacks for model variables (categorical VRS model) period 2004-2006

	Full sample (n=856)		Pure secondary school sample (n=611)		Cost efficiency model		Academic efficiency model		Academic effectiveness model		Overall efficiency model	
					FS[101] (n=856)	PS[102] (n=611)	FS (n=856)	PS (n=611)	FS (n=856)	PS (n=611)	FS (n=856)	PS (n=611)
	Mean	Std Dev.	Mean	Std Dev.	# Units[103] with non-zero slack[104] (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)	# Units with non-zero slack (%)
Teachers' salaries	3,652,524	1,966,280	4,255,695	1,973,906	0	0					0	0
Other expenditure	1,220,281	695,792	1,417,712	702,143	0	0					0	0
Roll	820	527	976	533	0	0	0	0				
# NCEA L1 passes	138	97	162	101			87 (10.2%)	58 (9.5%)	0	0		
# NCEA L2 passes	110	87	130	94			285 (33.3%)	232 (38%)	0	0		
# NCEA L3 passes	55	54	65	59			400 (46.8%)	293 (48%)	0	0		
# UE passes	52	51	60	56					8 (0.9%)	489 (0.8%)	44 (5.1%)	34 (5.6%)
# NCEA L3 M&E	14	20	16	22					458 (53.5%)	394 (64.5%)	563 (65.8%)	418 (68.4%)

[101] FS denotes the full sample.
[102] PS represents the pure secondary school sample.
[103] Ibid.
[104] Zero slack is at 0.01 level.

Appendix XI Statistics of non-zero weights for model variables (categorical VRS)

Table A11.1 Statistics of non-zero weights for model variables (categorical VRS) period 2009-2011

	Cost efficiency model						**Academic efficiency model**						**Academic effectiveness model**						**Overall efficiency model**					
	Full sample (N=865)			**Pure secondary schools (N=615)**			**Full sample (N=865)**			**Pure secondary schools (N=615)**			**Full sample (N=865)**			**Pure secondary schools (N=615)**			**Full sample (N=865)**			**Pure secondary schools (N=615)**		
	Mean	Std Dev	# Units[105] with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)
Teachers' salaries	5.85E-7[106]	5.98E-6	848 (98%)	2.78E-7	2.09E-7	600 (97.6%)													8.09E-7	1.6E-6	778 (89.9%)	6.64E-7	1.74E-6	532 (86.5%)
Other expenditures	4.22E-7	3.09E-6	818 (94.6%)	2.49E-7	2.54E-7	589 (95.8%)													6.72E-7	2.35E-6	679 (88.5%)	4.76E-7	1.41E-6	527 (85.7)
Roll	1.78E-3	1.20E-3	865 (100%)	1.39E-3	8.39E-4	615 (100%)	2.50E-3	4.44E-3	859 (99.3%)	1.72E-3	1.52E-3	608 (98.9%)												
# NCEA L1 passes							6.02E-3	5.03E-3	793 (91.7%)	4.20E-3	4.20E-3	519 (84.4%)	4.27E-3	0.018	378 (43.7%)	2.24E-3	6.70E-3	228 (37.1%)						
# NCEA L2 passes							3.16E-3	5.27E-3	605 (69.9%)	3.25E-3	5.17E-3	399 (64.9%)	3E-3	0.011	302 (35.3%)	2.04E-3	6.09E-3	193 (31.4%)						
# NCEA L3 passes							2.03E-3	6.37E-3	479 (55.4%)	1.45E-3	4.08E-3	323 (52.5%)	0.033	0.073	829 (95.8%)	0.028	0.043	595 (96.7%)						
# UE passes													0.031	0.055	853 (98.6%)	0.026	0.038	605 (98.4%)	0.034	0.054	832 (96.2%)	0.025	0.036	572 (93%)
# NCEA L3 M&E													0.021	0.040	646 (74.7%)	0.019	0.053	453 (73.7%)	0.014	0.062	187 (21.6%)	0.021	0.082	164 (26.7%)

105 Ibid.

106 5.85E-7=5.85×10^{-7}

Table A11.2 Statistics of non-zero weights for model variables (categorical VRS) period 2004-2006

	Cost efficiency model						Academic efficiency model						Academic effective model						Overall efficiency model					
	Full sample (N=856)			Pure secondary schools (N=611)			Full sample (N=856)			Pure secondary schools (N=611)			Full sample (N=856)			Pure secondary schools (N=611)			Full sample (N=856)			Pure secondary schools (N=611)		
	Mean	Std Dev	# Units[107] with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)	Mean	Std Dev	# Units with non-zero weights (%)
Teachers' salaries	4.35E-7	3.97E-7	839 (98%)	3.08E-7	3.67E-7	591 (96.7%)													1.05E-6	7.88E-6	703 (82.1%)	5.76E-7	1.35E-6	496 (81.2%)
Other expenditures	4.52E-7	5.85E-7	810 (94.6%)	3.83E-7	4.57E-7	579 (94.8%)													2.44E-6	1.26E-5	736 (86%)	1.01E-6	1.85E-6	547 (89.5%)
Roll	1.84E-3	1.28E-3	856 (100%)	1.39E-3	8.50E-4	611 (100%)	2.60E-3	4.75E-3	849 (99.2%)	1.84E-3	2.39E-3	604 (98.9%)												
# NCEA L1 passes							6.88E-3	5.87E-3	748 (87.4%)	5.68E-3	4.96E-3	534 (87.4%)	2.51E-3	7.64E-3	452 (52.8%)	2.43E-3	0.010	359 (58.8%)						
# NCEA L2 passes							3.22E-3	7.00E-3	544 (63.6%)	2.09E-3	4.75E-3	355 (58.1%)	0.005	0.029	654 (76.4%)	4.14E-3	0.021	466 (76.3%)						
# NCEA L3 passes							3.84E-3	8.45E-3	433 (50.6%)	3.36E-3	6.54E-3	304 (49.8%)	0.049	0.105	834 (97.4%)	0.035	0.094	588 (96.2%)						
# UE passes													0.051	0.099	842 (98.4%)	0.036	0.048	599 (98%)	0.047	0.098	800 (93.5%)	0.033	0.048	572 (93.6%)
# NCEA L3 M&E													0.014	0.039	366 (42.8%)	0.014	0.051	183 (30%)	0.031	0.110	272 (31.8%)	0.024	0.093	179 (29.3%)

[107] Ibid.

Appendix XII Both samples: percentage of schools with productivity changes

Table A12.1 Both samples: percentage of schools with increase/decrease in productivity for cost efficiency

Full sample: cost efficiency							
		Efficiency change		Frontier shift		Overall change	
SES[108]	N	+ (%)	- (%)	+ (%)	- (%)	+ (%)	- (%)
1	10	30	60	20	80	20	80
2	25	32	64	0	100	4	96
3	25	56	44	0	100	12	88
4	31	45.2	54.8	0	100	6.5	93.5
5	33	24.2	72.7	3	97	6.1	93.9
6	41	26.8	73.2	9.8	90.2	12.2	87.8
7	29	17.2	82.8	10.3	89.7	10.3	89.7
8	30	33.3	66.7	30	70	26.7	73.3
9	27	22.2	77.8	7.4	92.6	11.1	88.9
10	26	15.4	84.6	0	100	3.8	96.2
Total	277	30	69	7.6	92.4	10.8	89.2
Pure secondary school sample: cost efficiency							
		Efficiency change		Frontier shift		Overall change	
SES[109]	N	+ (%)	- (%)	+ (%)	- (%)	+ (%)	- (%)
1	6	33.3	50	100	0	66.7	33.3
2	21	28.6	66.7	0	100	4.8	95.2
3	20	65	35	0	100	25	75
4	22	50	50	0	100	13.6	86.4
5	24	33.3	66.7	0	100	16.7	83.3
6	29	27.6	72.4	0	100	3.4	96.6
7	24	20.8	75	0	100	8.3	91.7
8	18	16.7	83.3	0	100	0	100
9	21	38.1	61.9	0	100	14.3	85.7
10	14	21.4	78.6	0	100	0	100
Total	199	33.7	64.8	3	97	11.6	88.4

[108] Ibid.
[109] Ibid.

Table A12.2 Both samples: percentage of schools with increase/decrease in productivity for academic efficiency

Full sample: academic efficiency							
		Efficiency change		**Frontier shift**		**Overall change**	
SES[110]	**N**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**
1	10	60	30	100	0	100	0
2	25	68	24	100	0	88	12
3	25	56	44	100	0	88	12
4	31	74.2	25.8	100	0	83.9	16.1
5	33	69.7	27.3	97	3	81.8	18.2
6	41	51.2	46.3	100	0	82.9	17.1
7	29	55.2	44.8	100	0	79.3	20.7
8	30	56.7	43.3	93.3	6.7	66.7	33.3
9	27	51.9	48.1	88.9	11.1	70.4	29.6
10	26	65.4	30.8	34.6	65.4	65.4	34.6
Total	277	60.6	37.2	91.7	8.3	79.1	20.9
Pure secondary school sample: academic efficiency							
		Efficiency change		**Frontier shift**		**Overall change**	
SES[111]	**N**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**
1	6	50	33.3	100	0	66.7	33.3
2	21	52.4	38.1	100	0	81	19
3	20	60	40	95	5	80	20
4	22	68.2	31.8	81.8	18.2	77.3	22.7
5	24	66.7	29.2	83.3	16.7	83.3	16.7
6	29	69	31	72.4	27.6	79.3	20.7
7	24	62.5	37.5	83.3	16.7	70.8	29.2
8	18	66.7	27.8	38.9	61.1	55.6	44.4
9	21	57.1	33.3	19	81	47.6	52.4
10	14	64.3	28.6	0	100	71.4	28.6
Total	199	62.8	33.2	68.3	31.7	69.3	30.7

[110] Ibid.
[111] Ibid.

Table A12.3 Both samples: percentage of schools with increase/decrease in productivity for academic effectiveness

Full sample: academic effectiveness							
		Efficiency change		Frontier shift		Overall change	
SES[112]	N	+ (%)	- (%)	+ (%)	- (%)	+ (%)	- (%)
1	10	50	40	40	60	70	30
2	25	56	36	16	84	40	60
3	25	56	44	20	80	40	60
4	31	64.5	29	41.9	58.1	61.3	38.7
5	33	45.5	48.5	42.4	57.6	42.4	57.6
6	41	63.4	34.1	36.6	63.4	58.5	41.5
7	29	41.4	55.2	34.5	65.5	55.2	44.8
8	30	56.7	43.3	53.3	46.7	70	30
9	27	59.3	33.3	63	37	70.4	29.6
10	26	57.7	30.8	80.8	19.2	84.6	15.4
Total	277	55.6	39.4	43	57	58.5	41.5
Pure secondary school sample: academic effectiveness							
		Efficiency change		Frontier shift		Overall change	
SES[113]	N	+ (%)	- (%)	+ (%)	- (%)	+ (%)	- (%)
1	6	33.3	33.3	83.3	16.7	83.3	16.7
2	21	57.1	33.3	14.3	85.7	38.1	61.9
3	20	65	35	15	85	45	55
4	22	54.5	36.4	18.2	81.8	50	50
5	24	37.5	50	54.2	45.8	41.7	58.3
6	29	58.6	41.4	44.8	55.2	51.7	48.3
7	24	37.5	62.5	58.3	41.7	50	50
8	18	55.6	44.4	77.8	22.2	77.8	22.2
9	21	42.9	47.6	81	19	71.4	28.6
10	14	50	35.7	85.7	14.3	85.7	14.3
Total	199	50.3	43.2	49.2	50.8	55.8	44.2

[112] Ibid.
[113] Ibid.

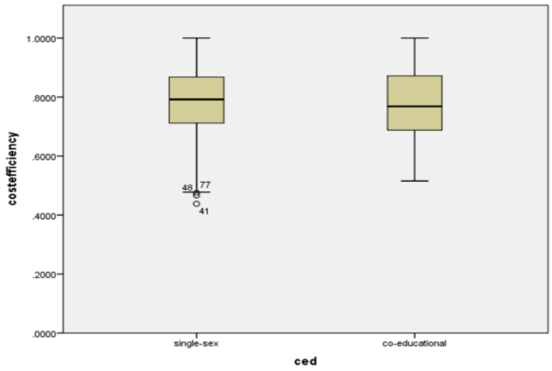
Table A12.4 Both samples: percentage of schools with increase/decrease in productivity for overall efficiency

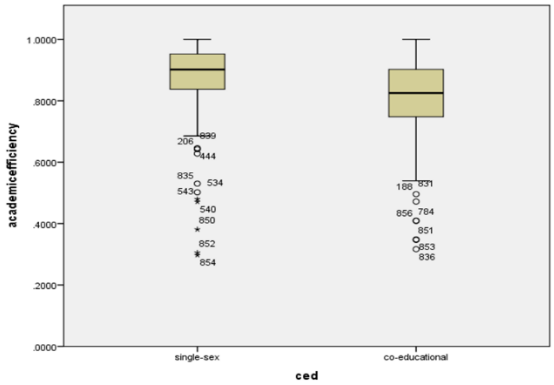
Full sample: overall efficiency							
		Efficiency change		**Frontier shift**		**Overall change**	
SES[114]	**N**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**
1	10	50	40	100	0	90	10
2	25	20	80	96	4	76	24
3	25	52	48	100	0	76	24
4	31	61.3	32.3	100	0	87.1	12.9
5	33	51.5	45.5	100	0	75.8	24.2
6	41	63.4	36.6	100	0	82.9	17.1
7	29	44.8	51.7	100	0	82.8	17.2
8	30	53.3	43.3	100	0	90	10
9	27	40.7	59.3	100	0	85.2	14.8
10	26	61.5	30.8	100	0	76.9	23.1
Total	277	50.9	46.2	99.6	0.4	81.9	18.1
Pure secondary school sample: overall efficiency							
		Efficiency change		**Frontier shift**		**Overall change**	
SES[115]	**N**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**	**+ (%)**	**- (%)**
1	6	33.3	50	100	0	83.3	16.7
2	21	23.8	76.2	95.2	4.8	71.4	28.6
3	20	60	40	95	5	75	25
4	22	59.1	36.4	100	0	72.7	27.3
5	24	62.5	37.5	100	0	70.8	29.2
6	29	69	31	100	0	75.9	24.1
7	24	62.5	37.5	100	0	75	25
8	18	61.1	38.9	100	0	77.8	22.2
9	21	47.6	47.6	100	0	85.7	14.3
10	14	57.1	28.6	92.9	7.1	71.4	28.6
Total	199	55.8	41.7	98.5	1.5	75.4	24.6

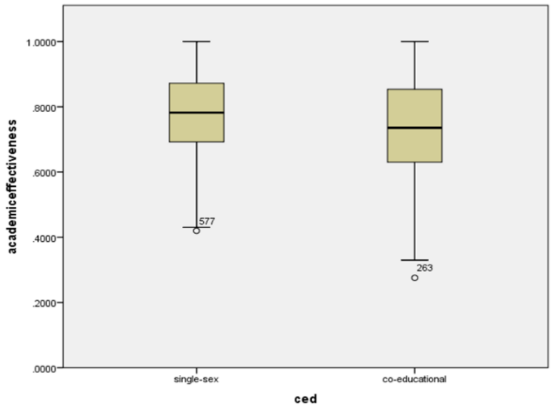
[114] Ibid.
[115] Ibid.

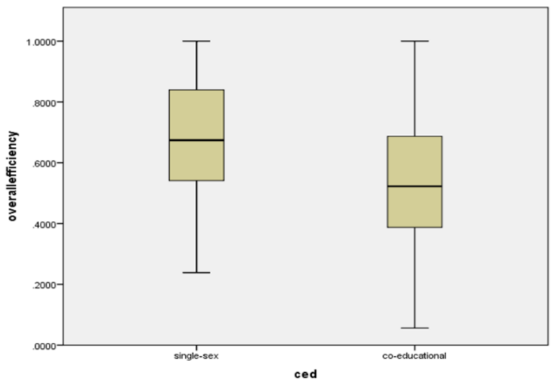
Appendix XIII Comparison between single-sex and co-educational schools

Figure A13.1 Comparison of efficiencies between single-sex and co-educational schools (full sample 2009-2011)









Appendix XIV Mäori pass rate vs. other ethnic groups pass rate

Figure A14.1 Full sample: comparison of NCEA Pass Rate between Mäori and Other Ethnic Groups for 2009-2011

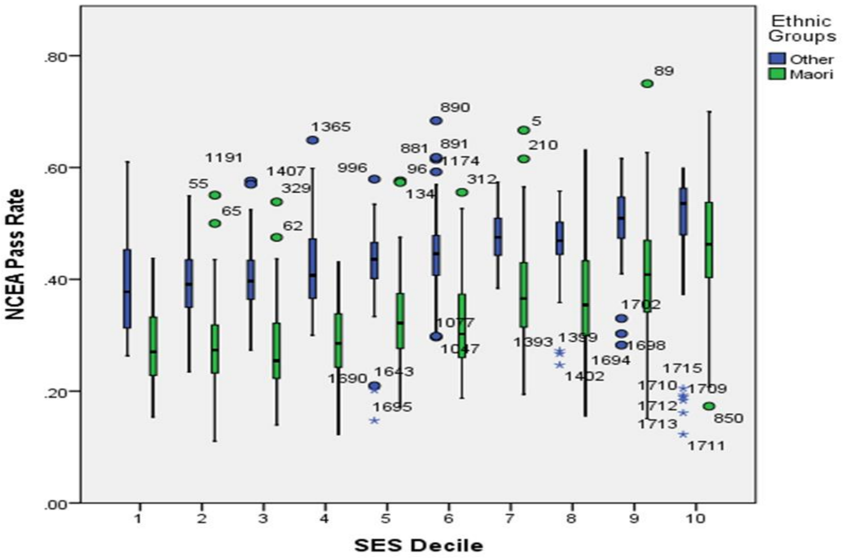


Figure A14.2 Full sample: comparison of UE Pass Rate between Mäori and Other Ethnic Groups for 2009-2011

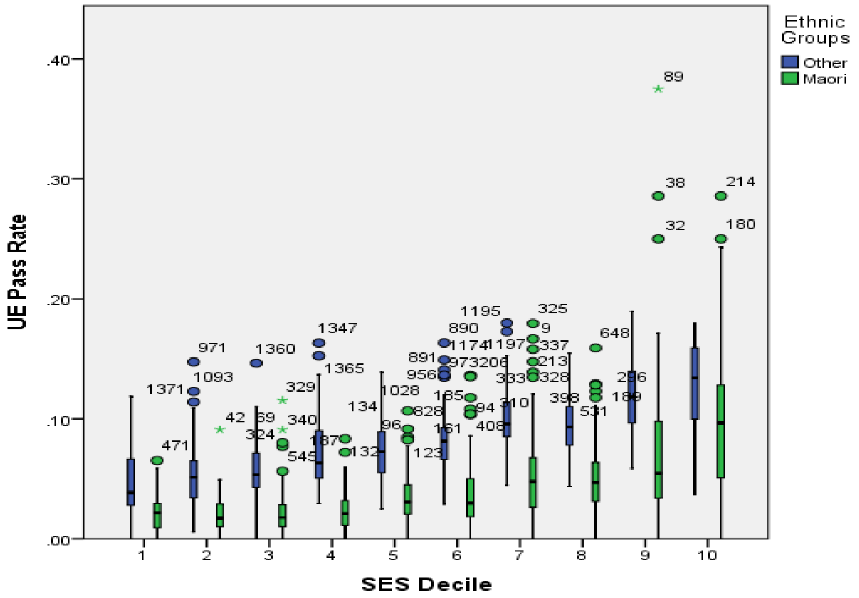


Figure A14.3 Full sample: comparison of NCEA Pass Rate between Mäori and Other Ethnic Groups for 2004-2006

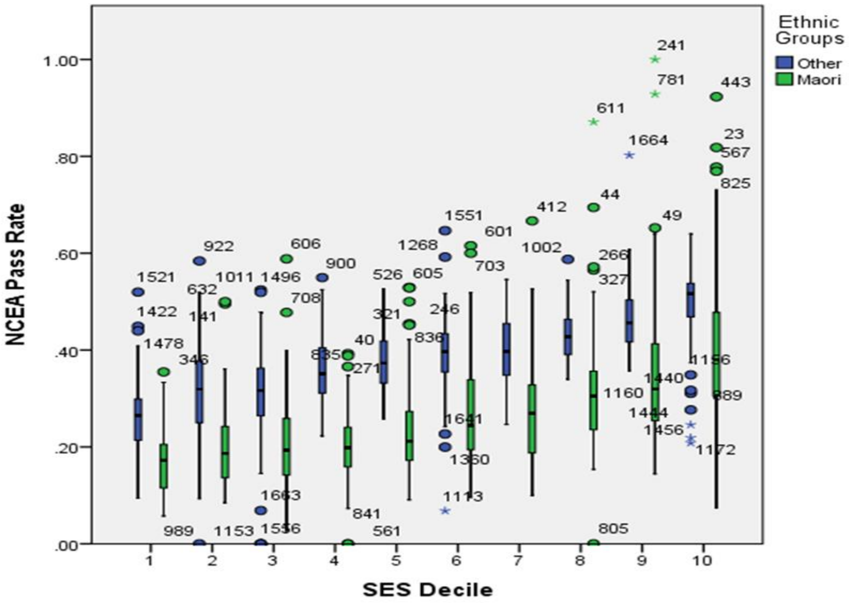
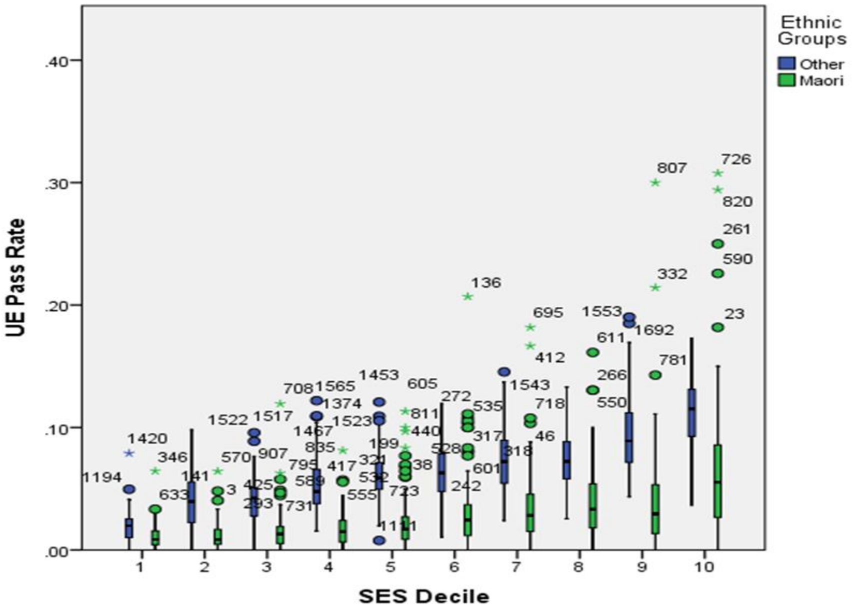


Figure A14.4 Full sample: comparison of UE Pass Rate between Mäori and Other Ethnic Groups for 2004-2006



Appendix XV Top ten schools in two test periods

Table A15.1 Top ten schools in two test periods

Test period:2009-2011								
ID	**Decile**	**Roll (2011/2006)**	**Type**	**Ownership**	**Gender**	**Year of ERO report**	**Years for next review**	**Mäori%**
TL1	10	3,016	Secondary	Non-integrated	Co-ed	2010	4-5	5%
TL2	10	974	Expanded	Integrated	Girls	2011	4-5	4%
TL3	7	826	Expanded	Integrated	Girls	2011	4-5	10%
TL4	3	706	Expanded	Integrated	Co-ed	2011	3	6%
TL5	9	2,196	Secondary	Non-integrated	Girls	2014	4-5	6%
TL6	4	2,233	Secondary	Non-integrated	Co-ed	2010	4-5	5%
TL7	1	1,444	Secondary	Non-integrated	Co-ed	2012	3	14%
TL8	3	728	Expanded	Non-integrated	Co-ed	2011	3	27%
TL9	3	184	Secondary	Integrated	Girls	2011	3	100%
TL10	10	1,272	Secondary	Non-integrated	Girls	2013	4-5	8%
Test period:2004-2006								
TE1	10	3049	Secondary	Non-integrated	Co-ed	2007	3	3%
TE2	10	956	Expanded	Integrated	Girls	2007	3	4%
TE3	3	733	Expanded	Integrated	Co-ed	2008	3	6%
TE4	10	1904	Secondary	Non-integrated	Girls	2006	3	5%
TE5	4	2284	Secondary	Non-integrated	Co-ed	2007	3	4%
TE6	1	1457	Secondary	Non-integrated	Co-ed	2006	3	12%
TE7	1	629	Secondary	Integrated	Girls	2007	3	5%
TE8	2	754	Secondary	Non-integrated	Girls	2008	3	36%
TE9	9	2605	Secondary	Non-integrated	Co-ed	2006	3	4%
TE10	8	1545	Secondary	Non-integrated	Co-ed	2007	3	8%

Appendix XVI Bottom ten schools in two test periods

Table A16.1 Bottom ten schools in two test periods

ID	Decile	Roll (2011/2006)	Type	Ownership	Gender	Year of ERO report	Years for next review	Mäori%
Test period: 2009-2011								
BL1	3	312	Secondary	Non-integrated	Co-ed	2011	1-2	72%
BL2	2	451	Secondary	Non-integrated	Co-ed	2012	1-2	56%
BL3	3	321	Secondary	Non-integrated	Co-ed	2012	3	60%
BL4	5	610	Secondary	Non-integrated	Co-ed	2012	3	29%
BL5	2	473	Secondary	Non-integrated	Co-ed	2011	3	63%
BL6	6	698	Secondary	Non-integrated	Co-ed	2012	1-2	18%
BL7	8	540	Expanded	Integrated	Boys	2012	3	14%
BL8	8	372	Expanded	Non-integrated	Co-ed	2011	3	12%
BL9	7	454	Secondary	Non-integrated	Co-ed	2011	3	16%
BL10	3	396	Expanded	Non-integrated	Co-ed	2012	1-2	41%
Test period:2004-2006								
BE1	3	276	Expanded	Integrated	Boys	2008	1	11%
BE2	2	363	Secondary	Non-integrated	Co-ed	2009	1-2	66%
BE3	7	472	Secondary	Non-integrated	Co-ed	2009	3	29%
BE4	5	276	Secondary	Non-integrated	Co-ed	2008	3	38%
BE5	4	296	Secondary	Non-integrated	Co-ed	2008	3	40%
BE6	1	351	Secondary	Non-integrated	Co-ed	2007	3	40%
BE7	7	703	Secondary	Non-integrated	Co-ed	2008	3	18%
BE8	2	761	Secondary	Non-integrated	Co-ed	2008	1	33%
BE9	4	375	Expanded	Non-integrated	Co-ed	2008	3	19%
BE10	6	396	Expanded	Non-integrated	Co-ed	2008	3	10%