

Combining the assumptions of variable and constant returns to scale in the efficiency evaluation of secondary schools

Victor V. Podinovski (corresponding author)

Warwick Business School, University of Warwick, Coventry CV4 7AL, UK

victor.podinovski@wbs.ac.uk

Phone: +44 (0)24 7652 8220; Fax: +44 (0)24 7652 4539

Ihsan Ismail

Educational Planning and Research Division, Ministry of Education, Putrajaya, 62604, Malaysia

Tatiana Bouzdine-Chameeva

KEDGE Business School, 680 cours de la Libération, 33405 Talence Cedex, France

Wenjuan Zhang

Warwick Business School, University of Warwick, Coventry CV4 7AL, UK

Keywords: data envelopment analysis; secondary education; schools; hybrid returns to scale; Malmquist index

Abstract

Our paper reports on the use of data envelopment analysis (DEA) for the assessment of performance of secondary schools in Malaysia during the implementation of the policy of teaching and learning mathematics and science subjects in the English language (PPSMI). The novelty of our application is that it makes use of the hybrid returns-to-scale (HRS) DEA model. This combines the assumption of constant returns to scale with respect to quantity inputs and outputs (teaching provision and students) and variable returns to scale (VRS) with respect to quality factors (attainment levels on entry and exit) and socio-economic status of student families. We argue that the HRS model is a better-informed model than the conventional VRS model in the described application. Because the HRS technology is larger than the VRS technology, the new model provides a tangibly better discrimination on efficiency than could be obtained by the VRS model. To assess the productivity change of secondary schools over the years surrounding the introduction of the PPSMI policy, we adapt the Malmquist productivity index and its decomposition to the case of HRS model.

1. Introduction

Data envelopment analysis (DEA) is an established methodology used for the assessment of efficiency and performance of organisations (Cooper et al., 2006; Thanassoulis et al., 2008). In the area of education, DEA has been widely used for the assessment of efficiency of the school provision at different levels (Grosskopf et al., 1999; Portela and Thanassoulis, 2001; Färe et al., 2006; Portela et al., 2012), universities and their departments (Avkiran, 2001; Thanassoulis et al., 2011), and the impact of education policies (Bradley et al., 2001; Grosskopf and Moutray, 2001).

The two conventional models traditionally employed in DEA studies, including applications in education, are based on the variable and constant returns to scale (VRS and CRS) technologies. Even if the true (best practice) technology is assumed to be VRS, the reference, or benchmark, CRS technology is often used as a part of the scale efficiency calculations. For the assessment of a particular education policy, it is common to use the Malmquist productivity index based on observations collected over a period of time, with its subsequent decompositions into different components (Johnes, 2008; Ouellette and Vierstraete, 2010; Thanassoulis et al., 2011).

The purpose of this paper is to demonstrate that in some applications, an example of which is the focus of our study, *a better-informed model* of the education technology may be obtained by the combination of VRS and CRS characteristics in one single formulation. To put our argument in the school education context, suppose that we have both quantity and quality performance factors (inputs and outputs). The quantity factors would typically include teachers (or teaching time) and students, while the quality factors may include a measure of academic attainment on entry (input) and on exit (output). It is often a legal requirement and an accepted managerial practice that there should be a certain ratio between teachers and students – this may vary between different schools or school types. Therefore, for example, a 10% increase of the student numbers requires the same increase of the teaching time. This indicates that the relationship between teachers and students may be assumed to be of the CRS type. It may, however, be difficult to argue that the CRS assumption extends to the quality inputs and outputs, as there may be no simple proportional relationship between these factors, in line with the quantity factors.

The above scenario poses a dilemma. In the described setting, the use of the CRS model for the underlying education technology is clearly unsubstantiated. On the other hand, the VRS technology is too conservative and does not use the information that students and

teachers satisfy the assumption of CRS. The latter implies that the VRS model underestimates the true technology and its use would lead to an overestimation bias of the efficiency scores.

In this paper we resolve the above dilemma by using the hybrid returns-to-scale (HRS) technology developed by Podinovski (2004a). The HRS technology is a convex polyhedral technology that exhibits CRS with respect to a subset of inputs and outputs (in the above scenario, with respect to teachers and students), and only VRS with respect to the remaining factors (quality inputs and outputs).

Our application concerns the performance of secondary schools in Malaysia in the provision of mathematics and science subjects. In 2003, the government of Malaysia initiated a new policy of teaching and learning these subjects in the English language (PPSMI). The policy was introduced in stages at different levels of school education and became a subject of wide debate within the Malaysian society (Ting, 2010).

In our study we focus on the upper secondary level of the school education and consider four cohorts of student. Two of these graduated in 2005 and 2006, and were educated before the PPSMI policy was introduced. The other two graduated in 2007 and 2008, and were educated under the new policy. Our data set includes 221 schools from three selected states in Malaysia that have complete data on the four cohorts.

This application is, to the best of our knowledge, the first reported application of the HRS technology with real-life data. Several contributions to the theory and methodology of DEA should be highlighted.

First, we extend the theoretical foundations of the HRS model by developing its new formulation. This is more intuitive than the original model and helps us to discuss the properties of the model and the assumptions behind it, in particular, in the context of school education. The original model is less intuitive but it is linear, and for this reason is used in the actual computations.

Second, we argue that the HRS model correctly represents our knowledge of the education technology, namely, that the quantity inputs and outputs satisfy the assumption of CRS, while the quality factors should be excluded. It is therefore a better-informed model than the standard VRS technology, in the described setting.

Third, our computations show that the HRS model produces a tangible practical improvement of the discrimination on efficiency scores, compared to the VRS model. This is because the HRS technology is an extension to the VRS technology and is, therefore, larger than the latter. Interestingly, the efficiency scores are also usually lower in the HRS model than in the CRS model (and the discrimination of the HRS model higher than that of CRS)

although the opposite relationship is observed for some individual schools in our sample. This is consistent with the theory and is explored in detail below.

Fourth, in order to assess the productivity change of the secondary education sector in the years surrounding the implementation of the PPSMI policy, we use the Malmquist productivity index and its decomposition. Although the extension of the Malmquist index to the case of HRS technologies is straightforward in principle, there is a particular difficulty that needs to be overcome. The conventional Malmquist index decomposition requires the assessment of efficiency of productive units in the VRS production technology and its reference CRS technology, constructed for the given and reference years. In these computations, the CRS technology is the cone extension of the VRS technology – it serves as the reference in the assessment of scale efficiency change and boundary shift. If, as in our study, the underlying technology is HRS, its cone extension is not the CRS but the cone-HRS (C-HRS) technology developed by Podinovski (2009). The latter is generally larger than both the HRS and standard CRS technology (and their union). Our application demonstrates how the Malmquist index can be decomposed and interpreted using the C-HRS technology.

The results of our study are largely consistent with the current discussion of the implications of the PPSMI policy on the school education (Ting, 2010). In particular, we find evidence that the average productivity of secondary schools in the teaching of mathematics and science subjects declined in the two years following the introduction of the policy. The biggest drop in school productivity was observed in 2007 – the first year of implementation of the policy at the upper secondary level, followed by a tangible, but not full, recovery in the year 2008. Our analysis also shows that this decline in schools in rural locations was greater than in urban locations. Moreover, in the case of rural locations, it had already started before the implementation of the policy but was less pronounced than in the years after its implementation. This indicates that the decline of productivity over all these years might have had a more complex nature than previously thought.

2. Preliminaries

2.1. Application background

The secondary education in Malaysia spans a period of five years: three years at the lower secondary level and two years at the upper secondary level. At the end of the lower secondary level students' performance is evaluated through an examination referred to as Lower Secondary Assessment (PMR). This examination is partly school based and adheres to the

national guidelines set by the Malaysian Examination Syndicate. Following this examination, students move to more specialised fields of study at the upper secondary level. At the end of the upper secondary level all students take the Malaysia Certificate of Education (SPM) examination. This examination is centrally administered and is considered to be equivalent to GCSE qualifications in England. A further one or two years of post-secondary education is required for entry to higher education.

Until 2003, all mathematics and science subjects at the secondary school level had been taught in the national language Bahasa Melayu. Starting in 2003, the Ministry of Education Malaysia introduced a new policy known as the Teaching and Learning of Mathematics and Science Subjects in English (PPSMI). The policy aimed at improving the English language skills of Malaysian students in the areas important for science, technology and international trade (Ting, 2010). Under the PPSMI policy, English became the medium of instruction in the teaching of all mathematics and science subjects. The implementation of this policy was carried out in stages beginning with the 2003 school year. At the upper secondary level the PPSMI policy was fully implemented starting in 2007.

Since its introduction, the purpose and benefits of the PPSMI policy became a subject of wide debate within the Malaysia society (Yahaya et al., 2009; Ting, 2010). Critics of the policy pointed, among other negative effects, poor students' performance in science subjects because of their weakness in the English language, and a disproportionate negative effect of the policy on schools in rural locations. The supporters of the policy highlighted the benefits of the policy and argued that the critics made unsubstantiated generalisations from limited studies of the effects of the policy.

Following the growing public pressure to reverse the PPSMI policy, in 2010 the Ministry of Education announced its decision to reinstate the national language for the teaching of mathematics and science subjects, starting in 2012. The main reason for this decision was the steady decline in the performance of students in these subjects as shown by the Ministry's own records and various studies (Ting, 2010).

2.2. Data collection

Data collection for this study started with an application to Economic Planning Unit (EPU), Prime Minister's Department Malaysia. Permission to conduct the research was granted by EPU in April 2008. With this permission, data for our study were obtained from Malaysian Examination Syndicate, Educational Planning and Research Division and School Division.

The Ministry of Education of Malaysia gave their permission to evaluate the performance of schools in the three northern states: Kedah, Penang, and Perlis. Overall, this included 303 secondary schools.

In order to assess the performance of schools before and after the implementation of the PPSMI policy, our analysis involved four cohorts of student. Two of these (2003 – 2005 and 2004 – 2006) were educated at the upper secondary level and took their PMR and SPM exams before the introduction of the policy. The other two cohorts, 2005 – 2007 and 2006 – 2008, were educated and took their exams entirely under the new policy.

Out of all schools in the three states only 237 had complete data on the four cohorts. This sample was reduced further by 16 schools that caused the well-known but rarely reported infeasibility problem in the application of the Malmquist index methodology (Färe et al., 1994b; Briec and Kerstens, 2009). Our final sample consisted of 221 schools from the three states.

3. Inputs and outputs

The focus of our study is on the performance of schools in Malaysia in the teaching of mathematics and science subjects at the upper secondary level. The choice of inputs and outputs for this assessment reflects this focus, and includes relevant quantity, quality and socio-economic factors. These are shown in Table 1.

< Table 1 here >

3.1. Inputs

The first five inputs represent the teaching provision in mathematics and science subjects. Based on the available data, these inputs are measured as the number of classes provided by a school in a specific subject. The class size typically varies between 25 and 30 students but on some occasions may be outside this range. In the Malaysian education system each class is supported by 1.5 teachers. Multiplying the number of classes in a subject by 1.5 would produce an accurate estimate of the number of teachers involved – but technically would be just a data rescaling exercise and result in an equivalent model. Our data set also showed the number of teachers employed by the school. We decided not to use this data because teachers often teach several different subjects and may have an uneven teaching load. It would therefore be unreliable to attribute a particular number of teachers to a certain science subject based on this type of data.

Two further inputs represent the number of students who joined the upper secondary level mathematics and science subjects with good grades (grades A and B in the relevant subjects in PMR examinations). These inputs control for the quality of the student intake and should positively influence the quality of the output.

The last input represents the number of students from families of high socio-economic status (SES). The latter is measured based on the family's income.

3.2. Outputs

The first five outputs represent the number of students on exit in each of the mathematics and science subjects. This measure represents a quantity (volume) output that requires resources (teaching time or classes) to produce. The authors are aware that some studies of school efficiency treat the number of students as input (Johnes et al., 2012). Because our goal is to model the education technology (and improve it by expanding the VRS model to HRS, as explained below), we differentiate between the inputs and outputs according to the theoretical axioms underpinning the definition of technology. The axiom of free disposability classes students as output because it is possible (within the technology) to teach fewer students rather than more, with the same resources.

The last five outputs measure the number of students who achieved good grades (grades A and B) in the SPM examination in a relevant subject on exit from school. These represent the quality of the education process.

4. Motivation of the new methodology

The VRS and CRS models are two conventional choices of DEA model. We would argue that the assumption of CRS is problematic in our application. Indeed, if all the eight inputs in Table 1 are increased by, for example, 10%, only the quantity outputs may be safely assumed to increase in the same proportion, as required by the CRS model. The five quality outputs may not necessarily increase in the same proportion because there is no simple linear relationship between the socio-economic status and attainment on entry as inputs, and attainment on exit as output.

The above implies that the VRS model might be a reasonable choice for our task. Furthermore, because we have a relatively large data set, we would expect the efficiency scores to be sufficiently discriminating and meaningful.

Let us show that the VRS model underutilizes our knowledge of the education process, and a larger model incorporating this knowledge would be a better estimate of the true technology. Consider the two scenarios.

4.1. Expansion scenario

Assume that the teaching resources of a school are increased by some factor $\beta > 1$ while keeping the quality and socio-economic inputs unchanged. (For example, $\beta = 1.1$ corresponds to an increase of 10%.) In our model specification this means that the number of classes in mathematics and science subjects increase by β . Based on the existing staffing policy, it should be possible for the school to increase the number of students in all the subjects by the same factor β . The above proportional relationship involves only the quantity inputs and outputs, and excludes quality and socio-economic factors.

As already discussed, the incorporation of the number of good students on exit in the above proportion between the quantity measures is questionable. Furthermore, we do not need to include the good students on entry and students from high SES in this proportion either because this is not needed for the increase of the quantity outputs by β .¹

Finally, if the number of classes and students are increased in proportion β , we do not know how the number of good students on exit might change, especially taking into account that the number of good students on entry is assumed to be constant. It may be argued that the number of good students may increase (for example, because the attainment on exit is not only a function of the attainment on entry, and therefore some students not classed as good on entry should become good students on exit). However, because we do not have any reliable information about this process, we assume *the worst-case scenario*: that the number of good students on exit stays constant for all $\beta > 1$.

The above can be summarized as the following assumption.

Assumption 1. For any school, it is technologically possible to change its inputs and outputs as follows:

- The number of classes in each of the mathematics and science subjects, and the corresponding number of students in each subject are increased by the same factor $\beta > 1$.

¹ By excluding the quality and SES inputs from the proportion, we state a stronger assumption and make the model of technology larger. Indeed, this assumption implies that it is possible to increase the number of students by β by increasing the number of classes by β , without an increase of quality and SES inputs.

- The number of good students on entry and exit, and the number of students from high SES remain unchanged.

4.2. Contraction scenario

Assume that the teaching resources of a school are reduced by some factor $\alpha \in [0,1]$ while keeping the quality and socio-economic inputs unchanged. (For example, $\alpha = 0.9$ corresponds to a reduction of 10%.) Similar to the expansion scenario, this would imply a reduction of the number of students in all mathematics and science subjects by the same factor α .

Similar to the expansion scenario, we rely on *the worst-case scenario* to describe the response of good students on exit to the reduction of classes and students with $\alpha < 1$. Note that we do not really know how the reduction of the number of classes might affect the number of students with good grades on exit. Therefore the only safe and undisputed assumption is that these numbers drop to zero, for even a small reduction of the inputs.

Obviously, this statement appears over-pessimistic and unrealistic, and the exact change of the number of good students on exit can be subject to further debate. Fortunately, this is not important *if the technology is convex*, such as the hybrid returns-to-scale technology formulated below. Theorem 1 proved and discussed in Podinovski (2004a) implies that, in our example, the described worst-case scenario is equivalent to the assumption of proportional change of the quality outputs by the same factor α , in line with the classes and student numbers. In other words, the assumption of convexity has a smoothing effect on the sudden drop of the outputs in the worst-case scenario.²

The following assumption describes the above worst-case scenario taking into account that the resulting technology is supposed to be convex.

Assumption 2. For any school, it is technologically possible to change its inputs and outputs as follows:

² The following simple example illustrates the reasons of this equivalence. Consider the DMU $A = (1,1)$, where the first component is input and the second is output. If we assume the logic of the worst-case scenario and change the input to any $0 \leq x < 1$, the output drops to zero and the resulting unit becomes $(x,0)$. When $x = 0$, we have the unit $B = (0,0)$. Let $0 \leq \alpha \leq 1$. If the technology is assumed convex, the combined unit $\alpha A + (1-\alpha)B = (\alpha, \alpha)$ is an element of this technology. Note that the unit (α, α) is obtained from A by the proportional reduction of its input and output by the factor α . Therefore, even though we assume that the output drops to zero with the reduction of the input by factor α , the assumption of convexity implies that a smoother decline of the output in proportion α is also feasible in the technology.

- The number of classes in each of the mathematics and science subjects, and the corresponding number of students in each subject are reduced by the same factor $\alpha \in [0,1]$.
- The number of good students on entry, and the number of students from high SES remain unchanged.
- The number of good students on exit is reduced by the same factor α . (As discussed, we can alternatively assume that the number of good students on exit is reduced to zero, or to any number within the range $[0,\alpha]$. The resulting model of technology would be the same.)

4.3. Further discussion of Assumptions 1 and 2

Although Assumptions 1 and 2 may seem to be too conservative, they are easy to justify and do not require any specific knowledge as to how the number of good students on exit actually responds to the change of the number of classes and students. Note that even these weak assumptions are stronger than those that define the VRS technology: the latter does not allow any scaling of the classes and students.

The treatment of the expansion and contraction scenarios in Assumptions 1 and 2 is symmetrical and based on the same principle. Namely, in both scenarios we assume the worst possible (and most defensible) outcome. In the case of expansion, this implies that the number of good students on exit remains constant, and in the case of contraction it is reduced to zero.

The worst-case statement of the contraction scenario is clearly unacceptable from the managerial point of view. For this reason we presented Assumption 2 in the more acceptable form stating that the number of such students decreases in proportion α , which is an equivalent assumption leading to the same model of technology.

Note that, even if we assume that the number of good students on exit drops to zero, Assumption 2 *does not state* what the actual reduction is. It only affirms that it is *technologically possible* to have no good students on exit – even if this cannot be allowed. It does not imply that a more reasonable (smoother) reduction of the number of good students on exit is impossible. As noted, the smoother proportional reduction is added to the model of technology by the convexity assumption.

A parallel may be drawn with the definition of conventional CRS technology based on the axioms stated in Banker et al. (1984). One of these is the assumption that the CRS

technology is convex. In particular, the latter implies that convex combinations of observed units are technologically possible. The sum of weights of the units used in a convex combination is equal to one. This does not contradict the fact that in the CRS technology the sum of weights can be any nonnegative number – the latter follows from *another assumption* stating that the units can be scaled up and down. Similarly, in our case Assumption 2 states that worst-case scenario is technologically possible – this does not contradict the fact that the convexity assumption adds more realistic units.

Assumptions 1 and 2 are the statements of *selective proportionality* (Podinovski 2004a) between a subset of inputs (classes) and a subset of outputs (student numbers). This is different from the assumption of full proportionality between all inputs and outputs in the CRS model. Below we describe the model that allows us to incorporate Assumptions 1 and 2 in the analysis.

5. The hybrid returns-to-scale model

5.1. Basic definitions and notation

The hybrid returns-to-scale (HRS) model was developed by Podinovski (2004a). The underlying HRS technology is an extension to the standard VRS model in which a subset of outputs is assumed proportional to a subset of inputs, as in the above Assumptions 1 and 2.

For a brief introduction of this technology, consider the following notation. Let T be a production technology with the set $I = \{1, \dots, m\}$ of inputs and the set $O = \{1, \dots, s\}$ of outputs. Its elements $(X, Y) \in T$ are referred to as DMUs. Let $J = \{1, \dots, n\}$ be the set of observed DMUs. These are stated as (X_j, Y_j) , $j \in J$.³

Let $I^P \subseteq I$ and $O^P \subseteq O$ denote the subsets of inputs and outputs that are assumed to be mutually proportional. In our application, the set I^P includes five quantity inputs – the number of classes in each of the five subjects. The set O^P includes five quantity outputs – the number of students in the corresponding subjects.

Let $I^{NP} = I \setminus I^P$ and $O^{NP} = O \setminus O^P$ be the complementary subsets of inputs and outputs, respectively. In our application, the set I^{NP} includes the number of good students on

³ A usual assumption (required for the correct statement of the technology and DEA models based on it) is made that each observed unit has at least one positive input and output. It is also assumed that, for each input $i \in I$ there exists an observed unit $j \in J$ such that its input $X_{ji} > 0$. Similarly, it is assumed that for each output $r \in O$, there exists an observed unit $j \in J$ such that its output $Y_{jr} > 0$.

entry in mathematics and science, and the number of students from high SES. The set O^{NP} includes five quality outputs: the number of good students on exit in each of the five subjects.

Any unit in technology T can be stated in the extended form as $(X, Y) = (X^P, X^{NP}, Y^P, Y^{NP})$, where the component subvectors X^P , X^{NP} , Y^P , and Y^{NP} correspond to the subsets I^P , I^{NP} , O^P , and O^{NP} , respectively.

5.2. Axioms

Following Podinovski (2004a), the HRS technology is defined by the following six axioms.

Axiom 1 (*Feasibility of observed data*). $(X_j, Y_j) \in T$, for all $j \in J$.

Axiom 2 (*Free disposability of inputs and outputs*). Let $(X, Y) \in T$. Consider any unit $(X', Y') \in \mathbb{R}_+^{m+s}$ such that $X' \geq X$ and $Y' \leq Y$. Then $(X', Y') \in T$.

Axiom 3 (*Convexity*). Technology T is a convex set.

Axiom 4 (*Selective proportionality: expansion scenario*). Let $(X, Y) \in T$. Then the unit $(X(\beta), Y(\beta)) = (\beta X^P, X^{NP}, \beta Y^P, Y^{NP}) \in T$, for any $\beta > 1$.

Axiom 5 (*Selective proportionality: contraction scenario*). Let $(X, Y) \in T$. Then the unit $(X(\alpha), Y(\alpha)) = (\alpha X^P, X^{NP}, \alpha Y^P, \mathbf{0}) \in T$, for any $\alpha \in [0, 1]$.

Axiom 6 (*Closedness*). Technology T is a closed set.

The first three of the above axioms are the standard axioms that define the VRS technology of Banker et al. (1984). Axioms 4 and 5 add DMUs to the technology by allowing selective proportional scaling of the inputs and outputs in the sets I^P and O^P . Note that these two axioms reflect the worst-case scenarios stated by Assumptions 1 and 2 above.

As proved in Podinovski (2004a, Theorem 1) if Axiom 3 is assumed, then Axiom 5 is equivalent to its following variant, which supports our discussion of the worst-case scenario:

Axiom 5*. Let $(X, Y) \in T$. Then the unit $(X(\alpha), Y(\alpha)) = (\alpha X^P, X^{NP}, \alpha Y^P, \alpha Y^{NP}) \in T$, for any $\alpha \in [0, 1]$.

Finally, Axiom 6 states that all limit points belong to the technology, that is T is a closed set. While this axiom is redundant in the VRS technology (it follows from Axioms 1 – 3 if only these axioms are assumed), it is no longer redundant and needs stating if the selective proportionality is allowed. The latter was demonstrated by an example in Podinovski (2004a).

5.3. The HRS technology

Our main development employs the *minimum extrapolation principle* that was first used in DEA for the definition of the VRS technology by Banker et al. (1984). To emphasize the importance of this principle, note that the VRS technology is conventionally assumed to be defined by the above Axioms 1 – 3. Note that there are many other technologies that satisfy the same three axioms: for example, the CRS and non-increasing returns-to-scale technologies. It is easy to prove that the VRS technology is a subset of any other such technology and is, therefore, the intersection of all technologies that satisfy Axioms 1 – 3. This fact implies that the VRS technology includes only those units that are necessary to satisfy the stated axioms. Any other technology includes some units that are arbitrary additions and do not follow from the axioms. For example, the CRS technology includes scaled units that are not assumed feasible by the stated Axioms 1 – 3. In other words, the VRS technology is the smallest technology that satisfies Axioms 1 – 3.

The definition of the HRS technology T_{HRS} by Podinovski (2004a) follows the same minimum extrapolation principle as in the case of VRS. Namely, the HRS technology does not only satisfy Axioms 1 – 6 but is also the smallest among all technologies that satisfies these axioms.

As a new development, below we define the related technology T^* based on Axioms 1 – 5 only. Technology T^* is a subset of technology T_{HRS} and allows a simpler and more intuitive interpretation than T_{HRS} . According to Theorem 3 proved below, T_{HRS} is “close enough” to T^* and is obtained by adding the “missing” limit (boundary) points to the latter⁴.

Definition 1. Technology T^* is the intersection of all technologies $T \subset \mathbb{R}_+^{m+s}$ that satisfy Axioms 1 – 5.

Definition 2. The HRS technology T_{HRS} is the intersection of all technologies $T \subset \mathbb{R}_+^{m+s}$ that satisfy Axioms 1 – 6.

The following is a new result that is useful for the interpretation of the above two technologies.

Theorem 1. Technology T^* is the set of all units $(X, Y) \in \mathbb{R}_+^{m+s}$ for which there exist vectors $\lambda, \alpha, \beta \in \mathbb{R}_+^n$ such that the following conditions are satisfied:

⁴ As shown in Podinovski (2004a), Axioms 1 – 5 do not imply Axiom 6. Therefore T^* is generally not a closed technology.

$$\sum_{j \in J} \lambda_j (\alpha_j \beta_j X_j^P) \leq X^P \quad (1.1)$$

$$\sum_{j \in J} \lambda_j (\alpha_j \beta_j Y_j^P) \geq Y^P \quad (1.2)$$

$$\sum_{j \in J} \lambda_j X_j^{NP} \leq X^{NP} \quad (1.3)$$

$$\sum_{j \in J} \lambda_j (\alpha_j Y_j^{NP}) \geq Y^{NP} \quad (1.4)$$

$$\alpha \leq \mathbf{1} \quad (1.5)$$

$$\beta \geq \mathbf{1} \quad (1.6)$$

$$\mathbf{1}^T \lambda = 1 \quad (1.7)$$

(In the above statement and below we use bold vectors $\mathbf{0}$ and $\mathbf{1}$ to denote vectors of zeros and ones, respectively. The dimension of these vectors is obvious from the context in which they are used. The superscript T in (1.7) denotes transposition.)

The meaning of conditions (1.1) – (1.7) is intuitively clear. Each term $j \in J$ on the left-hand side of inequalities (1.1) – (1.4) is obtained from the observed unit (X_j, Y_j) by first its selective proportional extension with $\beta_j \geq 1$ and its subsequent contraction with $\alpha \in [0, 1]$. According to assumed Axioms 4 and 5* (the latter equivalent to Axiom 5), the resulting units are considered to be in technology T^* . The whole left-hand side of conditions (1.1) – (1.4) is the convex combination of all such units taken with the weights λ_j . By Axiom 3 it represents a unit in technology T^* . Finally, the unit $(X, Y) = (X^P, X^{NP}, Y^P, Y^{NP})$ on the right-hand side is dominated by the unit on the left-hand side and is also in T^* .

The following result is the original statement of the HRS technology in the linear form.

Theorem 2 (Podinovski 2004a). Technology T_{HRS} is the set of all units $(X, Y) \in \mathbb{R}_+^{m+s}$ for which there exist vectors $\lambda, \mu, \nu \in \mathbb{R}_+^n$ such that the following conditions are satisfied:

$$\sum_{j \in J} (\lambda_j + \mu_j - \nu_j) X_j^P \leq X^P \quad (2.1)$$

$$\sum_{j \in J} (\lambda_j + \mu_j - \nu_j) Y_j^P \geq Y^P \quad (2.2)$$

$$\sum_{j \in J} \lambda_j X_j^{NP} \leq X^{NP} \quad (2.3)$$

$$\sum_{j \in J} (\lambda_j - \nu_j) Y_j^{NP} \geq Y^{NP} \quad (2.4)$$

$$\lambda - \nu \geq \mathbf{0} \quad (2.5)$$

$$\mathbf{1}^T \lambda = 1 \quad (2.6)$$

The relation between technologies T^* and T_{HRS} becomes apparent when we consider the following substitution of variables. Let some unit $(X, Y) \in T^*$. By Theorem 1, (X, Y) satisfies all conditions (1.1) – (1.7) with some vectors $\lambda, \alpha, \beta \in \mathbb{R}_+^n$. It is straightforward to verify that the unit (X, Y) then also satisfies conditions (2.1) – (2.6) with the same vector λ and the vectors $\mu, \nu \in \mathbb{R}_+^n$ defined from the following equations, for each j :

$$\lambda_j + \mu_j - \nu_j = \lambda_j \alpha_j \beta_j, \quad (3)$$

$$\lambda_j - \nu_j = \lambda_j \alpha_j. \quad (4)$$

Alternatively, equalities (3) and (4) are restated as

$$\mu_j = \lambda_j \alpha_j (\beta_j - 1), \quad (5)$$

$$\nu_j = \lambda_j (1 - \alpha_j). \quad (6)$$

Interestingly, the above substitution (5), (6) maps technology T^* on a subset of technology T_{HRS} but the converse is not true: a unit represented by conditions (2.1) – (2.6) may not always be represented by conditions (1.1) – (1.7). This result is not unexpected because technology T^* is generally a proper subset of T_{HRS} .

The following is a more precise statement of the above observation. We use notation $cl(T^*)$ for the closure of the set T^* . This is defined as the intersection of all closed sets T such that $T^* \subseteq T$.

Theorem 3. $T_{HRS} = cl(T^*)$.

Theorem 3 allows us to extend the above interpretation of units in technology T^* to technology T_{HRS} . Indeed, let $(X, Y) \in T_{HRS}$. Then one of the two cases is possible:

- 1) The unit (X, Y) is also in technology T^* . As discussed, any such unit is obtained by selective proportional expansion and/or contraction of one of the observed units, or is a convex combination of such scaled units, or is dominated by the latter.
- 2) The unit $(X, Y) \notin T^*$. By Theorem 3, (X, Y) is a limit point (unit) of some sequence of units $(X_t, Y_t) \in T^*$.

Therefore, in summary, any unit in the HRS technology T_{HRS} can be classed as one (or more than one) of the following five types: an observed unit, a selectively up-scaled or down-scaled unit, a convex combination of the former types, a dominated unit, or a limit unit of the previous types.

5.4. Efficiency assessment with the HRS model

The use of technology T_{HRS} in DEA models is straightforward. For example, the output radial efficiency of a particular DMU $(X_0, Y_0) \in T_{HRS}$ is obtained by inverting the optimal value of scalar variable η in the following program:

$$\max \eta, \text{ subject to } (X_0, \eta Y_0) \in T_{HRS}. \quad (7)$$

For practical computations, we use Theorem 2 to restate (7) as a linear program. Namely, the conditions of program (7) are stated in the form (2.1) – (2.6), in which the unit $(X, Y) = (X^P, X^{NP}, Y^P, Y^{NP})$ on the right-hand side is replaced by the unit $(X_0, \eta Y_0) = (X_0^P, X_0^{NP}, \eta Y_0^P, \eta Y_0^{NP})$.

Note that the conventional VRS technology is a subset of the HRS technology⁵, and therefore the efficiency discrimination of the HRS model (7) should be better than the VRS model. Our computations presented below support this general observation.

6. Computations and analysis

6.1. Analysis of efficiency

The first step of our analysis involved the assessment of output radial efficiency of all 221 schools in our sample, separately for each of the four years 2005 – 2008. For example, the efficiency of schools in year 2005 was assessed using the data for cohort 2003-2005. Similarly, the efficiency in year 2006 was evaluated based on the data for cohort 2004-2006.

We performed computations in the HRS technology described above. For reference purposes, we also computed the output radial efficiency of schools in the VRS and CRS technologies. All computations were programmed and executed in a commonly available linear optimiser. The computation results are summarised in Table 2.

< Table 2 here >

⁵ The description of VRS technology is obtained from conditions (2), where μ and ν are zero vectors.

As theoretically expected, the DEA model based on the HRS technology discriminates better on efficiency than the model based on the VRS technology, because the latter is a subset of the former. In our analysis the mean efficiency in the HRS model is between 4% to 5% lower than in the VRS model.

It is also interesting to compare the results of the HRS and CRS models. Although above we argued against the use of the CRS model with the inputs representing quality and socio-economic factors, theoretically it is still worth comparing the scores, especially because the CRS model is conventionally regarded as the best discriminating DEA model. Neither the HRS nor the CRS technology is a subset of the other, and there are schools in the sample whose efficiency is lower in the CRS model than in HRS, and vice versa. However, overall the HRS model outperforms the CRS model in each year of analysis – the mean efficiency in the former is between 2% to 4% lower than in the latter. An interesting implication of this is that, even if there are doubts about whether the quality inputs and outputs should be included in the proportionality assumption, the negative decision (resulting in the use of HRS over CRS) is beneficial in both ways: this improves our confidence in the model specification and improves the discrimination of the model. This contrasts with the usual dilemma involving the choice of VRS or CRS technology: the former is a safer assumption but the latter leads to a better discrimination.

The efficiencies in Table 2 are also interesting from the policy perspective. The lower mean efficiency implies a bigger gap between the best and worst performing schools. Although the difference between the years is not very large (for the same model), we can still observe that the gap is the biggest (by about a 1% to 2% margin) in year 2008, when the PPSMI policy was in the second year of implementation. This suggests that some schools have benefited more from the policy than the others.

6.2. *The Malmquist index*

Below we describe the application of the Malmquist index and its decomposition for the analysis of change of school productivity, observed over the implementation period of the PPSMI policy. The traditional use of the Malmquist index involves the VRS technology and its cone extension – the CRS technology (Färe et al., 1994a, 1994b).⁶ Because the technology in our application is assumed to be HRS (instead of VRS), we need its cone extension to

⁶ Alternative decompositions of the Malmquist index are discussed by Lovell (2003) and Grosskopf (2003).

obtain a similar decomposition to the conventional approach. The following known result describes the required technology which is referred to as the cone-HRS (C-HRS) technology.

Theorem 4 (Podinovski 2009). The cone extension T_{C-HRS} to technology T_{HRS} is the set of all units $(X, Y) \in \mathbb{R}_+^{m+s}$ for which there exist vectors $\lambda, \mu, \nu \in \mathbb{R}_+^n$ such that all conditions (2.1) – (2.5) are satisfied.

Theorem 4 implies that the cone technology T_{C-HRS} is obtained by omitting the normalising equation (2.6) from the statement of technology T_{HRS} .

Following traditional notation, let units (X^t, Y^t) and (X^{t+1}, Y^{t+1}) represent the same school in years t and $t+1$, respectively. Let $E_{HRS}^t(X^t, Y^t)$ and $E_{C-HRS}^t(X^t, Y^t)$ be the output radial efficiency of the given school in technologies T_{HRS}^t and T_{C-HRS}^t respectively, constructed by the data set from year t . Define $E_{C-HRS}^{t+1}(X^t, Y^t)$ and $E_{C-HRS}^t(X^{t+1}, Y^{t+1})$ in a similar way. For example, $E_{C-HRS}^{t+1}(X^t, Y^t)$ denotes the output radial efficiency of the given school (X^t, Y^t) relative to the frontier of technology T_{C-HRS}^{t+1} .

The Malmquist index of productivity change can be defined and interpreted on the cone technology T_{C-HRS} in the same way as on the cone CRS technology. By taking the geometric mean of the period t and period $t+1$ Malmquist productivity indexes, we obtain

$$M_{C-HRS}(X^t, Y^t, X^{t+1}, Y^{t+1}) = \left[\frac{E_{C-HRS}^t(X^{t+1}, Y^{t+1})}{E_{C-HRS}^t(X^t, Y^t)} \times \frac{E_{C-HRS}^{t+1}(X^{t+1}, Y^{t+1})}{E_{C-HRS}^{t+1}(X^t, Y^t)} \right]^{\frac{1}{2}}, \quad (8)$$

As in the conventional case of CRS, the meaning of (8) is straightforward. The first ratio in (8) assesses the change of the average productivity of the school using the cone technology in year t . The second term uses technology in year $t+1$.⁷

⁷ Let T be any technology that satisfies the very weak assumptions stated in Podinovski (2004b) – for example T can be a VRS or HRS technology. Let C be the cone extension of T (for example, CRS or C-HRS, respectively). Podinovski (2004b, page 233) proves that the radial efficiency (obviously the same in both input and output orientations) of any unit $(X, Y) \in T$ in technology C is equal to the minimum of the ratio δ/γ , subject to the condition $(\delta X, \gamma Y) \in T$. In other words, the output (or input) radial efficiency of the unit (X, Y) in C is equal to its average productivity expressed as a fraction of the average productivity of its most productive scale size (MPSS). For example, if $E_{C-HRS}^t(X^t, Y^t) = 0.5$, the unit's average productivity is 50% compared to its MPSS in the form $(\delta X, \gamma Y)$ found in technology T_{HRS}^t . If at the same time $E_{C-HRS}^t(X^{t+1}, Y^{t+1}) = 0.6$, the average productivity of the unit (compared in the latter case to a generally different MPSS but still in the same technology T_{HRS}^t) has improved in year $t+1$ compared to year t by a factor $0.6/0.5 = 1.2$.

Repeating the logic of the Malmquist index decomposition as in Färe et al. (1994a), we decompose the school productivity change (8) in two adjacent years t and $t+1$ as follows:

$$M_{C-HRS}(X^t, Y^t, X^{t+1}, Y^{t+1}) = TE\Delta \times SE\Delta \times T\Delta,$$

where

$$TE\Delta = \frac{E_{HRS}^{t+1}(X^{t+1}, Y^{t+1})}{E_{HRS}^t(X^t, Y^t)}, \quad (9)$$

$$SE\Delta = \frac{SE^{t+1}(X^{t+1}, Y^{t+1})}{SE^t(X^t, Y^t)} = \left(\frac{E_{C-HRS}^{t+1}(X^{t+1}, Y^{t+1})}{E_{HRS}^{t+1}(X^{t+1}, Y^{t+1})} \right) \bigg/ \left(\frac{E_{C-HRS}^t(X^t, Y^t)}{E_{HRS}^t(X^t, Y^t)} \right), \quad (10)$$

$$T\Delta = \left[\frac{E_{C-HRS}^t(X^{t+1}, Y^{t+1})}{E_{C-HRS}^{t+1}(X^{t+1}, Y^{t+1})} \times \frac{E_{C-HRS}^t(X^t, Y^t)}{E_{C-HRS}^{t+1}(X^t, Y^t)} \right]^{\frac{1}{2}}. \quad (11)$$

The above terms (9) – (11) can be interpreted in the conventional sense, with the substitution of VRS and CRS by, respectively, HRS and C-HRS technologies. Namely, the pure technical efficiency change $TE\Delta$ is the ratio of the output radial efficiency of the school in year $t+1$ to its efficiency in year t .

The scale efficiency change $SE\Delta$ is the ratio of the scale efficiencies of the school in years $t+1$ and t . Each of the two scale efficiencies is calculated as the ratio of the output radial efficiency of the school in the cone technology C-HRS to its output radial efficiency in the assumed technology HRS (Podinovski, 2009). This is similar to the calculations in the case of the VRS technology and its cone extension CRS.

The term ΔT represents the geometric mean of two terms. The first of these shows the change of the average productivity of school (X^{t+1}, Y^{t+1}) as the technology progresses from year t to $t+1$ (Grosskopf, 2003). The second has the same meaning for the school represented as (X^t, Y^t) . Overall, ΔT has a traditional interpretation as an indicator of technical change, or boundary shift.

6.3. Analysis of productivity change

We used the described Malmquist index methodology to analyse the productivity change of secondary schools in Malaysia over three years: 2006, 2007 and 2008, in

comparison to 2005 as the base year⁸. Table 3 shows the computational results. These are the geometrical averages of the corresponding terms computed for all 221 schools in the sample.

< Table 3 here >

Our results show that both the technical efficiency and scale efficiency of the schools in the sample have changed little in relation to the base year 2005. The only exception is the pre-policy year 2006 which exhibited an almost 2% increase in the technical efficiency.

The technical efficiency change in years 2006, 2007 and 2008 shown in Table 3 is consistent with the means of efficiency scores in the HRS model in the corresponding years as shown in Table 2. In particular, the mean HRS score is the highest in 2006, followed by 2007, 2005 and 2008, respectively. The geometric mean $TE\Delta$ of pure technical efficiency change in Table 2 points to the same order.⁹

In contrast, the technical change, or boundary shift, is much more noticeable. In all three years after 2005 the most productive schools (and consequently, the boundary of the cone C-HRS technology) in each year have exhibited a decline in productivity. It is interesting that this decline is observed in both pre-policy year 2006 (decline over 3%) and in the years 2007 and 2008 under the PPSMI policy (decline over 6% in 2007 and about 2.5% in 2008).

The overall Malmquist index is the product of the three above components, and its decline is primarily explained by the negative effect of the technology change.

< Table 4 here >

Furthermore, each school in our data set was classed as either urban or rural. In total, there were 108 urban and 113 rural schools. Table 4 shows the Malmquist index and its decomposition in the two different technologies: one based only on the urban, and the other on rural schools. It is easy to notice that the schools of both location types essentially replicate the pattern already observed in their combined analysis. Namely, as can be seen from the last column of Table 4, the overall productivity of schools has, regardless of their

⁸ The calculation of the Malmquist index in relation to the same base year 2005 follows the approach of Grifell-Tatjé and Lovell (1996). We also investigated the application of the Malmquist index to the consecutive pairs of years 2005-2006, 2006-2007 and 2007-2008. The results exhibited a strong non-circularity property which is a well-known theoretical fact (Fried et al., 2008). This made them less interesting for policy interpretation.

⁹ Note that Table 2 shows the arithmetic means of efficiency scores, while Table 3 shows the geometric means of their ratios. Therefore, the two types of mean, while indicating the same direction of change, produce close but generally different numerical estimates of the change.

location, deteriorated mostly in the first year of the implementation of the PPSMI policy (shown as year 2007), followed by the noticeable recovery in the following year 2008. The three multiplicative components of the Malmquist index shown in Table 4 indicate that most of this deterioration was due to the technical change, while the schools changed little with respect to their technical and scale efficiency relative to the retracting efficient frontier.

Judging by the overall Malmquist index and its main changing component (technical change), the productivity dynamics of rural schools lagged behind the productivity of urban schools in all three years of analysis. In particular, in 2007 the average productivity of a rural school declined by more than 7% compared to the base year 2005, while the decline for urban schools was more modest at approximately 4.5%. In 2008 these numbers changed to about 3.8% and 1.4% respectively. This might indicate that the initial implementation of the PPSMI policy was more challenging for rural schools than urban schools.

Overall, our results are consistent with the existing literature on the effects of the PPSMI policy (Ting, 2010) and indicate that:

- There was a decline in the average school productivity over the three years 2006 – 2008 compared to the base year 2005. The biggest drop was registered in 2007 – the first year of the PPSMI policy. However, this improved in the following year 2008. The fact that the decline was also observed in year 2006 before the new policy was implemented, suggests that the reasons for the deterioration of average school productivity might have had complex roots¹⁰.
- While the average school productivity in years 2006 – 2008 was below their productivity in the base year 2005, most schools retained their technical efficiency (proximity to the best practice HRS frontier in relative terms), even when the frontier was retracting from the highest levels observed in 2005.
- The observed productivity decline followed the same pattern for both rural and urban schools: the biggest drop in 2007 and signs of recovery in 2008. However, the decline was deeper in rural locations than in urban schools.

7. Conclusion

In this paper we reported an application of the hybrid returns-to-scale (HRS) model to the analysis of efficiency of secondary schools in Malaysia, and analysed the change of their

¹⁰ This finding is not unusual. For example, a decrease of the productivity in the higher education sector in England has been observed by Thanassoulis et al. (2011).

productivity over a period 2005 – 2008. During these years the new policy of teaching mathematics and science subjects in the English language (PPSMI) was implemented.

The HRS model exhibits CRS with respect to a subset of inputs and outputs while keeping the remaining inputs and outputs subject to the assumption of VRS. In this paper we argued that the HRS model is a better-informed model in the context of education technologies than the conventional VRS model. Indeed, it is common practice to link the number of students at school to the teaching resources. It is therefore reasonable to assume the CRS (proportional relationship) between the number of students in individual subjects to the teaching provision (number of classes in our application). The quality of the teaching process is represented in the model by the attainment metrics on entry and exit, and measures of socio-economic status. We argued that the quality inputs and outputs might not exhibit the same proportionality property and should therefore be excluded from the assumption of CRS.

Our data set included 221 secondary schools in the Malaysian states Kedah, Penang and Perlis. Our computations showed that the HRS model discriminated better on the efficiency scores than the conventional VRS model. The mean school efficiency in the HRS model was 4 – 5% lower than the mean efficiency in the VRS model, in each of the four years of analysis 2005 – 2008. Interestingly, the HRS model also outperformed the CRS model by 2 – 4% on the mean efficiency. (This comparison was performed only for academic reasons, as we argued against the actual use of CRS in the application).

We used the Malmquist index and its decomposition to analyse the productivity change of Malaysian secondary schools in the two years before the implementation of the PPSMI policy, and two years after. The conventional application of the Malmquist index is performed in the CRS technology which is the cone extension of the underlying VRS technology. Because in our application the underlying technology is HRS, its cone extension is not CRS but a larger set called cone-HRS (C-HRS) technology. In our paper we used an adaptation of the Malmquist index based on the HRS and C-HRS technologies.

The application of the Malmquist index methodology to our data set showed that in the three years 2006, 2007 and 2008 there was a decline in the average productivity of the upper second level education at Malaysian schools driven predominantly by the technical change. The decline was most pronounced in 2007 – the first year in which the PPSMI policy was implemented, although there was a tangible recovery in the following year. We also showed that the productivity of urban schools was affected to a smaller extent than rural schools, although both location types followed the same pattern.

Acknowledgements

The authors are thankful to the Education Planning and Research Division, Examination Syndicate, School Division, and the State Education Departments of Kedah, Perlis, and Penang under the Ministry of Education of Malaysia for their permission to collect and use the data for this study.

Appendix A: Proofs

Proof of Theorem 1. The proof consists of two parts stated in Lemmas 1 and 2.

Lemma 1. Technology T^* satisfies Axioms 1 – 5.

Proof of Lemma 1. T^* obviously satisfies Axioms 1 and 2. In particular, the unit (X', Y') in Axiom 2 satisfies (1.1) – (1.7) with the same vectors $\lambda, \alpha, \beta \in \mathbb{R}_+^n$ as the unit (X, Y) , and therefore, $(X', Y') \in T^*$. To prove that T^* satisfies Axiom 3, consider any two units (X^1, Y^1) and (X^2, Y^2) from T^* . Then these satisfy (1.1) – (1.7) with some vectors $\lambda^1, \alpha^1, \beta^1 \in \mathbb{R}_+^n$ and $\lambda^2, \alpha^2, \beta^2 \in \mathbb{R}_+^n$, respectively. Consider $\delta \in [0, 1]$ and define $(\tilde{X}, \tilde{Y}) = \delta(X^1, Y^1) + (1 - \delta)(X^2, Y^2)$. We need to prove that (\tilde{X}, \tilde{Y}) satisfies (1.1) – (1.7) with some vectors $\tilde{\lambda}, \tilde{\alpha}, \tilde{\beta} \in \mathbb{R}_+^n$.

Indeed, the required vectors are found from the following equations:

$$\tilde{\lambda} = \delta\lambda^1 + (1 - \delta)\lambda^2, \quad (12)$$

$$\tilde{\lambda}_j \tilde{\alpha}_j = \delta\lambda_j^1 \alpha_j^1 + (1 - \delta)\lambda_j^2 \alpha_j^2, \quad \forall j \in J, \quad (13)$$

$$\tilde{\lambda}_j \tilde{\alpha}_j \tilde{\beta}_j = \delta\lambda_j^1 \alpha_j^1 \beta_j^1 + (1 - \delta)\lambda_j^2 \alpha_j^2 \beta_j^2, \quad \forall j \in J. \quad (14)$$

Note that (12) fully defines vector $\tilde{\lambda}$. From (13), and taking into account (12) and (1.5), we have

$$\tilde{\alpha}_j = \frac{\delta\lambda_j^1 \alpha_j^1 + (1 - \delta)\lambda_j^2 \alpha_j^2}{\tilde{\lambda}_j} = \frac{(\delta\lambda_j^1) \alpha_j^1 + ((1 - \delta)\lambda_j^2) \alpha_j^2}{\delta\lambda_j^1 + (1 - \delta)\lambda_j^2} \in [0, 1], \quad \forall j \in J. \quad (15)$$

Similarly, from (14), and by (13) and (1.6), we have

$$\tilde{\beta}_j = \frac{\delta\lambda_j^1 \alpha_j^1 \beta_j^1 + (1 - \delta)\lambda_j^2 \alpha_j^2 \beta_j^2}{\tilde{\lambda}_j \tilde{\alpha}_j} = \frac{(\delta\lambda_j^1 \alpha_j^1) \beta_j^1 + ((1 - \delta)\lambda_j^2 \alpha_j^2) \beta_j^2}{\delta\lambda_j^1 \alpha_j^1 + (1 - \delta)\lambda_j^2 \alpha_j^2} \geq 1, \quad \forall j \in J. \quad (16)$$

Note that the derivation in (15) assumes that $\tilde{\lambda}_j \neq 0$. If $\tilde{\lambda}_j = 0$, (12) implies that both terms $\delta\lambda_j^1 = (1 - \delta)\lambda_j^2 = 0$, and (13) and (14) are satisfied if we arbitrarily take $\tilde{\alpha}_j = \tilde{\beta}_j = 1$. Similarly, in (16) we assume that $\tilde{\lambda}_j \tilde{\alpha}_j \neq 0$. If $\tilde{\lambda}_j \tilde{\alpha}_j = 0$, then either $\tilde{\lambda}_j = 0$ (this case was dealt with above), or $\tilde{\lambda}_j \neq 0$ but $\tilde{\alpha}_j = 0$. In the latter case both terms on the right-hand side of (13) are equal to zero, and (14) is satisfied if we arbitrarily take $\tilde{\beta}_j = 1$.

Note that (14) implies that (\tilde{X}, \tilde{Y}) satisfies (1.1) and (1.2) with vectors $\tilde{\lambda}, \tilde{\alpha}, \tilde{\beta}$. Similarly, (13) implies (1.4), and (12) implies (1.3) and (1.7). Finally, (15) and (16) imply (1.5) and (1.6). Therefore, T^* satisfies Axiom 3.

To prove that T^* satisfies Axiom 4, let unit (X, Y) satisfy (1.1) – (1.7) with some $\lambda, \alpha, \beta \in \mathbb{R}_+^n$. Consider any $\tilde{\beta} \geq 1$ and define the unit $(X(\tilde{\beta}), Y(\tilde{\beta})) = (\tilde{\beta}X^P, X^{NP}, \tilde{\beta}Y^P, Y^{NP})$ as in Axiom 4. It is straightforward to verify that $(X(\tilde{\beta}), Y(\tilde{\beta}))$ satisfies (1.1) – (1.7) with vectors λ , α and $\hat{\beta} = \beta\tilde{\beta}$. Similarly, consider any $\tilde{\alpha} \in [0, 1]$ and define $(X(\tilde{\alpha}), Y(\tilde{\alpha})) = (\tilde{\alpha}X^P, X^{NP}, \tilde{\alpha}Y^P, \tilde{\alpha}Y^{NP})$ as in Axiom 5* (equivalent to A5). Then $(X(\tilde{\alpha}), Y(\tilde{\alpha}))$ satisfies (1.1) – (1.7) with vectors λ , $\hat{\alpha} = \alpha\tilde{\alpha}$ and β . \square

Lemma 2. $T^* \subseteq T'$, where $T' \subset \mathbb{R}_+^{m+s}$ is any technology that satisfies Axioms 1 – 5.

Proof of Lemma 2. Let $(X, Y) \in T^*$, and therefore (X, Y) satisfies (1.1) – (1.7) with some $\lambda, \alpha, \beta \in \mathbb{R}_+^n$. Consider any $j \in J$ such that $\lambda_j > 0$. Because T' satisfies Axiom 1, $(X_j, Y_j) \in T'$. By Axiom 4, $(X_j(\beta_j), Y_j(\beta_j)) = (\beta_j X_j^P, X_j^{NP}, \beta_j Y_j^P, Y_j^{NP}) \in T'$. Furthermore, applying the selective contraction with $\alpha \in [0, 1]$ to the above unit $(X_j(\beta_j), Y_j(\beta_j))$, by Axiom 5* we conclude that $(\tilde{X}_j, \tilde{Y}_j) = (\alpha_j \beta_j X_j^P, X_j^{NP}, \alpha_j \beta_j Y_j^P, \alpha_j Y_j^{NP}) \in T'$. By Axiom 3, the unit on the left-hand side of (1.1) – (1.4) is in T' . This unit dominates the unit (X, Y) on the right-hand side. By Axiom 2, $(X, Y) \in T'$. \square

Proof of Theorem 3. Because T_{HRS} satisfies Axioms 1 – 5, $T^* \subseteq T_{HRS}$. Therefore, $cl(T^*) \subseteq cl(T_{HRS}) = T_{HRS}$. To prove the converse embedding $T_{HRS} \subseteq cl(T^*)$, by definition of T_{HRS} it suffices to prove that $cl(T^*)$ satisfies Axioms 1 – 6. Axioms 1 and 6 are obvious. Because T^* is convex, by Theorem 6.2 in Rockafellar (1970), $cl(T^*)$ is convex and satisfies Axiom 3.

To prove that T^* satisfies Axiom 2, consider any $(X, Y) \in cl(T^*)$ and let unit (X', Y') be such that $X' \geq X$ and $\mathbf{0} \leq Y' \leq Y$. We need to prove that $(X', Y') \in cl(T^*)$. Indeed, there exists a sequence of units $(X_t, Y_t) \in T^*$, $t = 1, 2, \dots$, such that

$$(X_t, Y_t) \rightarrow (X, Y) \text{ as } t \rightarrow +\infty. \quad (17)$$

Define the vector $(W_t, Z_t) = (X_t, Y_t) - (X, Y)$, $\forall t$. Then $(X_t, Y_t) = (X, Y) + (W_t, Z_t)$, $\forall t$, and $\|(W_t, Z_t)\| \rightarrow 0$ as $t \rightarrow +\infty$. Define $(X'_t, Y'_t) = (X', Y') + (W_t, Z_t)$. Then $X'_t = X' + W_t \geq X + W_t = X_t$ and, similarly, $Y'_t \leq Y_t$, $\forall t$. If all vectors $Y'_t \geq \mathbf{0}$, then by Axiom 2, $(X'_t, Y'_t) \in T^*$. Because $(X'_t, Y'_t) \rightarrow (X', Y')$ as $t \rightarrow +\infty$, by Axiom 6 $(X', Y') \in cl(T^*)$. If some vectors Y'_t have negative components, we change all such components to zero, for all such vectors, and repeat the above proof.

To prove that T^* satisfies Axiom 4, consider any unit $(X, Y) \in cl(T^*)$ and $\beta \geq 1$. We need to prove that the unit $(X(\beta), Y(\beta)) = (\beta X^P, X^{NP}, \beta Y^P, Y^{NP}) \in cl(T^*)$. Indeed, consider the sequence of units $(X_t, Y_t) \in T^*$ as in (17). Because T^* satisfies Axiom 4, the unit

$$(X_t(\beta), Y_t(\beta)) = (\beta X_t^P, X_t^{NP}, \beta Y_t^P, Y_t^{NP}) \in T^*, \forall t.$$

Because $(X_t(\beta), Y_t(\beta)) \rightarrow (X(\beta), Y(\beta))$ as $t \rightarrow +\infty$, $(X(\beta), Y(\beta)) \in cl(T^*)$.

The proof that T^* satisfies Axiom 5* (and therefore A5) is similar. □

References

- Avkiran, N.K., 2001. Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences* 35, 57 – 80.
- Banker, R.D., Charnes A., Cooper W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30, 1078 – 1092.
- Bradley, S., Johnes, G., Millington, J., 2001. The effect of competition on the efficiency of secondary schools in England. *European Journal of Operational Research* 135, 545 – 568.
- Briec, W., Kerstens, K., 2009. Infeasibilities and directional distance functions: with application to the determinateness of the Luenberger productivity indicator. *Journal of Optimization Theory and Applications* 141, 55 – 73.
- Cooper, W.W., Seiford, L.M., Tone, K., 2006. *Introduction to Data Envelopment Analysis and Its Uses*. Springer, New York.
- Färe, R., Grosskopf, S., Lindgren, B., Roos, P., 1994a. Productivity developments in Swedish hospitals: A Malmquist output index approach. In: Charnes, A., Cooper, W.W., Lewin, A.Y., Seiford, L.M. (Eds.), *Data Envelopment Analysis: Theory, Methods and Applications*. Kluwer Academic Publishers, Boston, MA, pp. 253 – 272.
- Färe, R., Grosskopf, S., Norris, M., Zhang, Z., 1994b. Productivity growth, technical progress, and efficiency change in industrialized countries, *American Economic Review* 84, 66 – 83.
- Färe, R., Grosskopf, S., Førsund, F.R., Hayes, K., Heshmati, A., 2006. Measurement of productivity and quality in non-marketable services: with application to schools. *Quality Assurance in Education* 14, 21 – 36.
- Fried, H.O., Lovell, C.A.K., Schmidt, S.S., 2008. Efficiency and productivity. In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (Eds.), *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, New York, pp. 3 – 91.
- Grifell-Tatjé, E., Lovell, C.A.K., 1996. Deregulation and productivity decline: the case of Spanish savings banks. *European Economic Review* 40, 1281 – 1303.
- Grosskopf, S., 2003. Some remarks on productivity and its decompositions. *Journal of Productivity Analysis* 20, 459 – 474.
- Grosskopf, S., Hayes, K.J., Taylor, L.L., Weber, W.L., 1999. Anticipating the consequences of school reform: a new use of DEA. *Management Science* 45, 608 – 620.

- Grosskopf, S., Moutray, C., 2001. Evaluating performance in Chicago public high schools in the wake of decentralization. *Economics and Education Review* 20, 1 – 14.
- Johnes, J., 2008. Efficiency and productivity change in the English higher education sector from 1996/97 to 2004/5. *The Manchester School* 76, 653 – 674.
- Johnes, J., Bradley, S., Little, A., 2012. Efficiency in the further education sector in England. *Open Journal of Statistics* 2, 131 – 140.
- Lovell, C.A.K., 2003. The decomposition of Malmquist productivity indexes. *Journal of Productivity Analysis* 20, 437 – 458.
- Ouellette, P., Vierstraete, V., 2010. Malmquist indexes with quasi-fixed inputs: an application to school districts in Québec. *Annals of Operations Research* 173, 57 – 76.
- Podinovski, V.V., 2004a. Bridging the gap between the constant and variable returns-to-scale models: selective proportionality in data envelopment analysis. *Journal of the Operational Research Society* 55, 265 – 276.
- Podinovski, V.V., 2004b. Efficiency and returns to scale on the “no free lunch” assumption only. *Journal of Productivity Analysis* 22, 227 – 257.
- Podinovski, V.V., 2009. Production technologies based on combined proportionality assumptions. *Journal of Productivity Analysis* 32, 21 – 26.
- Portela, M.C.S., Camanho, A.S., Borges, D., 2012. Performance assessment of secondary schools: the snapshot of a country taken by DEA. *Journal of the Operational Research Society* 63, 1098 – 1115.
- Portela, M.C.A.S., Thanassoulis, E., 2001. Decomposing school and school-type efficiency. *European Journal of Operational Research* 132, 357 – 373.
- Rockafellar, R.T., 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.
- Thanassoulis, E., Portela, M.C.S., Despić, O., 2008. Data envelopment analysis: the mathematical programming approach to efficiency analysis. In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (Eds.), *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, New York, pp. 251 – 420.
- Thanassoulis, E., Kortelainen, M., Johnes, G., Johnes, J., 2011. Costs and efficiency of higher education institutions in England: a DEA analysis, *Journal of the Operational Research Society* 62, 1282 – 1297.
- Ting, S.H., 2010. Impact of language planning on language choice in friendship and transaction domains in Sarawak, Malaysia. *Current Issues in Language Planning* 11, 397 – 412.

Yahaya, M.F.B., Noor, M.A.B.M., Mokhtar, A.A.B., Rawian, R.B.M., Othman, M.B., Jusoff, K., 2009. Teaching of mathematics and science in English: The teachers' voices. *English Language Teaching* 2, 141 – 147.

Table 1

Inputs and outputs

Inputs	Outputs
Mathematics classes	Mathematics students
Science classes	Science students
Physics classes	Physics students
Biology classes	Biology students
Chemistry classes	Chemistry students
Good mathematics students on entry	Good mathematics students on exit
Good science students on entry	Good science students on exit
Students from the high SES group	Good physics students on exit
	Good biology students on exit
	Good chemistry students on exit

Table 2

Descriptive statistics of efficiency scores in different DEA models.

Year	Model	Minimum	Maximum	Mean
2005	HRS	.59	1.00	.9016
	VRS	.62	1.00	.9475
	CRS	.62	1.00	.9345
2006	HRS	.50	1.00	.9134
	VRS	.64	1.00	.9551
	CRS	.53	1.00	.9352
2007	HRS	.57	1.00	.9105
	VRS	.61	1.00	.9570
	CRS	.61	1.00	.9399
2008	HRS	.58	1.00	.8929
	VRS	.60	1.00	.9527
	CRS	.60	1.00	.9392

Table 3

Malmquist index and its decomposition calculated on the base year 2005.

Year	Technical efficiency change $TE\Delta$	Scale efficiency change $SE\Delta$	Technical change $T\Delta$	Malmquist Index $TE\Delta \times SE\Delta \times T\Delta$
2006	1.0190	1.0002	0.9687	0.9873
2007	1.0073	0.9979	0.9372	0.9420
2008	0.9992	0.9971	0.9751	0.9715

Table 4

Malmquist indexes by location calculated on the base year 2005.

Location	Year	Technical efficiency change $TE\Delta$	Scale efficiency change $SE\Delta$	Technical change $T\Delta$	Malmquist Index $TE\Delta \times SE\Delta \times T\Delta$
Urban	2006	1.0281	0.9980	0.9749	1.0003
	2007	1.0121	0.9969	0.9469	0.9554
	2008	0.9993	1.0076	0.9794	0.9862
Rural	2006	1.0131	1.0018	0.9618	0.9762
	2007	1.0027	0.9988	0.9281	0.9296
	2008	0.9915	0.9993	0.9711	0.9622