

Jill Johnes

Pomiar efektywności*

STRESZCZENIE: Celem niniejszego rozdziału jest przede wszystkim identyfikacja i przedstawienie różnych metod pomiaru efektywności stosowanych w kontekście oceny funkcjonowania instytucji edukacyjnych w tym szkół wyższych. Ponadto dokonano przeglądu badań empirycznych wykorzystujących te metody na wszystkich poziomach edukacji.

SŁOWA KLUCZOWE: efektywność, szkolnictwo wyższe, edukacja, pomiar, metody nieparametryczne, DEA

1. Wstęp

W ostatnich latach coraz większym zainteresowaniem cieszą się badania związane z pomiarem funkcjonowania instytucji działających w sektorze edukacji. Dane, które odzwierciedlają różne aspekty funkcjonowania szkół i uniwersytetów, są coraz częściej regularnie publikowane w wielu krajach, zarówno na poziomie lokalnym, jak i krajowym. Dane te często wyrażone są w formie relacji przedstawiających efektywność uzyskiwanych rezultatów (np. stosunek uczniów, którzy uzyskali w danym roku minimalny poziom wymaganych osiągnięć do ogólnej liczby uczniów) lub efektywność wykorzystanych nakładów (np. koszt na studenta). W taki sposób łatwo

* Praca ukazała się oryginalnie jako: Johnes, J. (2004). Efficiency measurement. W: G. Johnes i J. Johnes (red.) *International Handbook on the Economics of Education* (613-742), Cheltenham–Northampton, MA: Edward Elgar Publishing Ltd. Dziękujemy autorce i wydawcy za zgodę na tłumaczenie. W nawiasach dla wybranych terminów podano także terminy i skróty w języku angielskim (przyp. tłum.).

jest stworzyć „tabele ligowe” szkół i uczelni, które będą wskazówką do dokonania wyborów oraz będą informować o alokacji zasobów.

Zagrożenia z używaniem takich wskaźników wydajności są obecnie dobrze znane. Po pierwsze, instytucje takie jak szkoły czy uniwersytety działają w różnych warunkach i w różnym środowisku, co nie jest należycie odzwierciedlone w prostych wskaźnikach. Po drugie, instytucje edukacyjne wytwarzają szereg rezultatów, zużywając wiele nakładów. Dlatego stosunek jednego wyniku do jednego nakładu nie obrazuje w pełni funkcjonowania organizacji, która prowadzi wieloraką działalność i może być stosowany tylko jako wskaźnik cząstkowy.

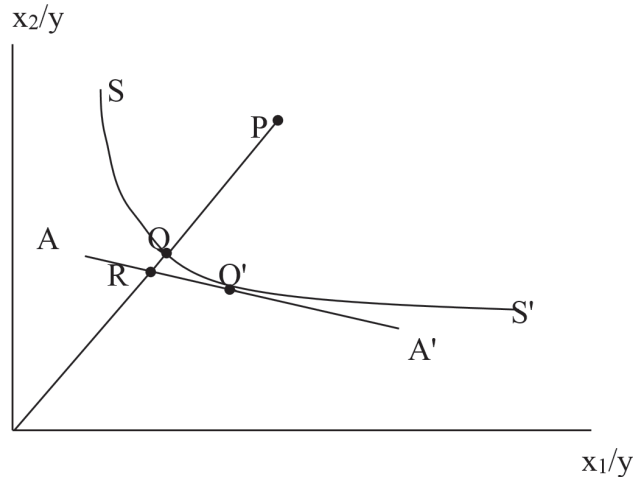
W konsekwencji opracowano wiele technik, które są stosowane w kontekście instytucji działających na rynku edukacji w celu zmierzenia rzeczywistej efektywności ich funkcjonowania. Narzędzia statystyczne przeszły od etapu estymacji za pomocą metody najmniejszych kwadratów (MNK) do analizy z użyciem stochastycznej metody granicznej, a proste wskaźniki wyrażone jako stosunek jednego wyniku do jednego nakładu zostały zastąpione poprzez złożone wskaźniki efektywności uzyskane dzięki metodzie programowania liniowego (ang. *linear programming* LP). Celem tego rozdziału jest przede wszystkim identyfikacja i przedstawienie różnych metod pomiaru efektywności stosowanych w kontekście edukacji oraz dokonanie przeglądu badań empirycznych wykorzystujących te metody na wszystkich poziomach edukacji w celu oceny ich użyteczności dla usługobiorców, menadżerów i decydentów. Rozdział składa się z sześciu części, a pierwszą z nich jest wstęp. W drugiej zaprezentowano niektóre z definicji efektywności wywodzące się z teorii produkcji. Części trzecia i czwarta zostały poświęcone metodom pomiaru efektywności, a piąta wadom stosowanych metod oraz przedstawieniu przykładów badań, w których zostały one wykorzystywane do oceny funkcjonowania instytucji edukacji. Wnioski przedstawiono w części szóstej.

2. Definicja efektywności

Przydatnym punktem wyjścia do dyskusji na temat pomiaru efektywności jest praca Farrella (1957) definiująca trzy typy efektywności, które zostaną przedstawione poniżej. Weźmy pod uwagę firmę, która używa nakładów x_1 i x_2 , aby wyprodukować wynik y w wielkości opisanej przez punkt P na Rysunku 1.1. W kontekście edukacji firmą może być szkoła, która produkuje absolwentów (y) zużywając nakłady: pracowników (x_1) oraz uczniów (x_2). Krzywa SS' przedstawia izokwantę łączącą wszystkie kombinacje x_1 i x_2 , z których efektywna firma może wybrać poziom produkcji przy zastosowaniu stałych korzyści skali (ang. *constant returns to scale* CRS). Efektywność techniczna (TE) firmy wytwarzającej w punkcie P jest zdefiniowana jako procent nakładów wykorzystywanych przez firmę, który jest niezbędny do wyprodukowania wyników w punkcie P , i można ją zapisać jako:

$$TE = \frac{OQ}{OP}.$$

Rysunek 1.1 Pomiar efektywności technicznej zorientowanej na nakłady



Wielkość $1-TE$ jest miarą nieefektywności technicznej i oznacza procent, o jaki mogłyby zostać zmniejszone nakłady bez zmniejszania poziomu produkcji w punkcie P (Worthington 2001). Kolejna miara efektywności może zostać zdefiniowana przez wprowadzenie cen nakładów. Na rysunku zostały one zaznaczone jako prosta AA' , która przedstawia stosunek cen dwóch nakładów¹. Punkt Q' reprezentuje zarówno efektywność techniczną, jak i efektywność alokacyjną (lub cenową), przedstawia ideę, że nakłady są używane w optymalnych (minimalizujących koszty) proporcjach przy danych cenach nakładów. Ponieważ koszty produkcji w punkcie Q' są także reprezentowane przez punkt R , który znajduje się na prostej AA' , miara efektywności alokacyjnej (ang. *allocative efficiency AE*) firmy jest zdefiniowana jako:

$$AE = \frac{OR}{OQ}.$$

Punkty Q i Q' są efektywnie technicznie, ale koszty produkcji w punkcie Q' są częścią OR/OQ kosztów w punkcie Q . Dlatego redukcja kosztów produkcji (przy danych cenach nakładów), która mogłaby zostać osiągnięta przy użyciu kosztowo efektywnych proporcji nakładów, jest prezentowana przez odcinek RQ . Wreszcie,

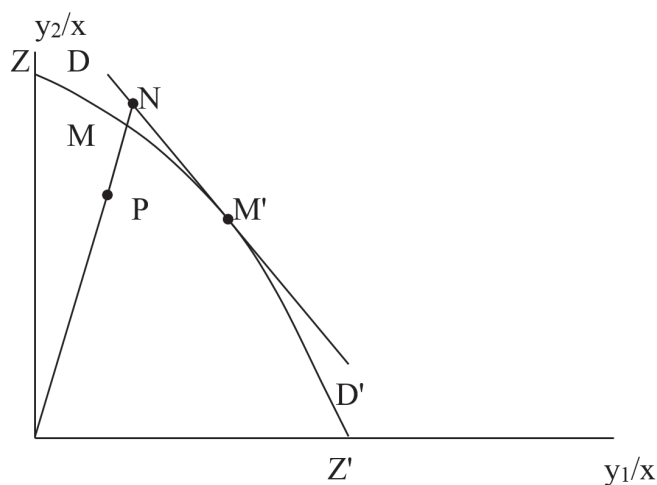
¹ Zakładamy, że ceny nakładów są dane, a firma nie działa w warunkach monopsonu.

efektywność całkowita (ekonomiczna) (ang. *overall efficiency OE*) danej firmy jest definiowana jako:

$$OE = \frac{OR}{OP}.$$

Efektywność całkowita jest wynikiem efektywności technicznej i alokacyjnej. Wielkość $1-OE$ jest miarą nieefektywności, a odcinek RP prezentuje zmniejszenie kosztów produkcji, które mogłyby zostać osiągnięte przez firmę przy przejściu z punktu P do punktu, gdzie następuje minimalizacja kosztów: Q' .

Rysunek 1.2 Pomiar efektywności technicznej zorientowanej na wyniki



Podejście przyjęte przez Farrella (1957) jest zorientowane na nakłady, ponieważ ma na celu odpowiedzieć na pytanie: o ile firma może proporcjonalnie zmniejszyć nakłady bez zmiany ilości wyprodukowanych wyników? Można także zastosować podejście zorientowane na wyniki, w którym szuka się odpowiedzi na pytanie: o ile można proporcjonalnie zwiększyć produkcję dla danej wielkości nakładów? Podejście nakierowane na wyniki jest zilustrowane na Rysunku 1.2. W tym wypadku firma produkuje dwa wyniki: y_1 i y_2 z danego nakładu x . Przykładowo może to być uniwersytet, którego wynikami są rezultaty działalności związanej z kształceniem studentów (y_1) oraz wyniki działalności naukowej (y_2), a nakładem jest czas pracy pracowników uczelni (x). Krzywa ZZ' jest krzywą możliwości produkcyjnych, a efektywność techniczna jest stosunkiem uzyskanej produkcji do tej, która jest maksymalnie możliwa do uzyskania:

$$TE = \frac{OP}{OM}.$$

W tym przypadku $1-TE$, który jest miarą nieefektywności, określa, o ile może wzrosnąć produkcja bez zwiększania poziomu nakładów obserwowanych w punkcie produkcji P . Efektywność alokacyjna jest wyznaczona przez określenie linii tych samych przychodów: DD^2 . Punkt M' jest zarówno efektywny technicznie, jak i alokacyjnie. Ponieważ w punkcie M' mamy takie same przychody jak w punkcie N wzdłuż linii DD' , efektywność alokacyjna może zostać zmierzona jako:

$$AE = \frac{OM}{ON}.$$

Punkty M i M' są efektywne technicznie, ale przychody z produkcji w punkcie N są stosunkiem OM/ON części przychodów w punkcie M' . Zwiększenie przychodów, które firma mogłaby osiągnąć poprzez produkcję właściwej (maksymalizującej przychody) proporcji wyników jest opisana jako odcinek MN . Efektywność całkowita jest zatem:

$$OE = \frac{OP}{ON},$$

i jest wynikiem efektywności technicznej i alokacyjnej. Zatem $1-OE$ to miara nieefektywności, a odległość PN wskazuje, o ile firma mogłaby zwiększyć przychody poprzez przesunięcie punktu produkcji z punktu P do punktu M' – maksymalizacji przychodów.

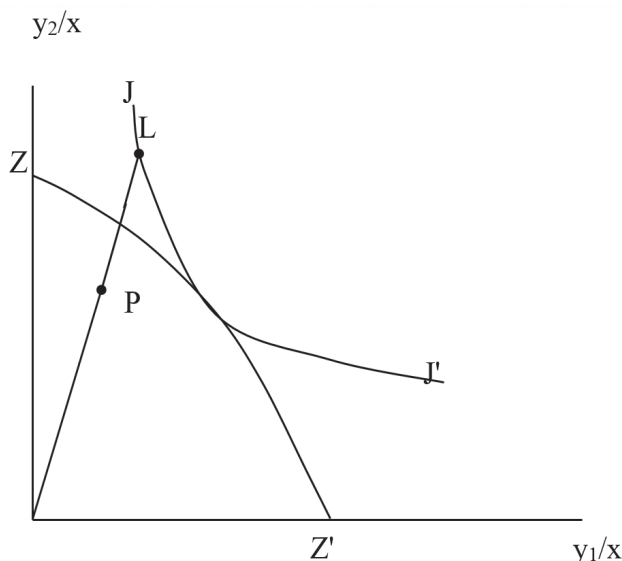
W przypadku stałych korzyści skali miary efektywności technicznej uzyskane przy orientacji na nakłady oraz przy orientacji na wyniki są takie same. Różnią się one jednak w warunkach zmiennych korzyści skali (ang. *variable returns to scale* VRS).

Miary efektywności określone za pomocą Rysunków 1.1 i 1.2 nie zawierają żadnej informacji na temat preferencji. W celu zdefiniowania efektywności społecznej (ang. *social efficiency* SE) należy dodać funkcję dobrobytu społecznego przeprowadzoną przez dwa wyniki. Dokonano tego na Rysunku 1.3, na którym dodano krzywą JJ' , która opisuje preferencje pomiędzy dwoma rezultatami. Idealny punkt produkcji z punktu widzenia dobrobytu społecznego znajduje się w punkcie L' . Jednak użyteczność przy zestawieniu wyników w punkcie L' jest równa użyteczności pochodzącej z produktów w punkcie L , więc miara efektywności społecznej dla firmy znajdującej się w punkcie P , która jest zdefiniowana jako stosunek obserwowanego produktu do poziomu produkcji oczekiwanej przez społeczeństwo, jest równa:

² Zakładamy, że ceny wyników są dane, a firma nie jest monopolistą.

$$SE = \frac{OP}{OL}.$$

Rysunek 1.3 Pomiar efektywności społecznej zorientowanej na wyniki



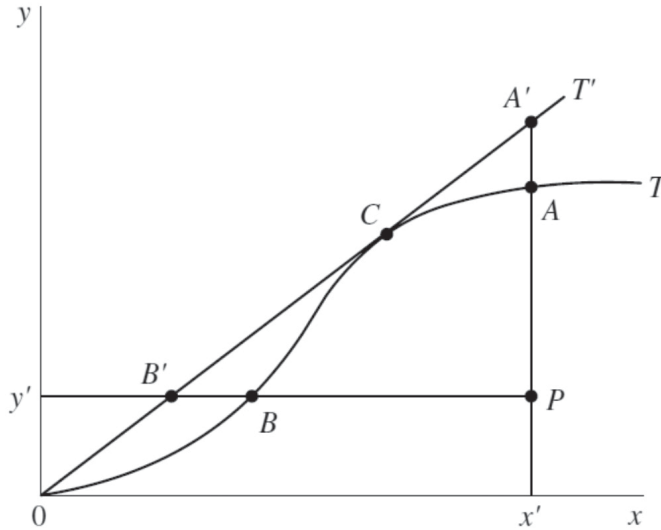
Zauważmy, że z tej definicji wynika, że możliwa jest efektywność społeczna bez efektywności alokacyjnej lub efektywność alokacyjna bez efektywności społecznej.

Wszystkie dotychczasowe definicje efektywności zostały oparte na założeniu stałych korzyści skali. Założenie to jest dość restrykcyjne, ponieważ firma może być nieefektywna z powodu nieodpowiedniej wielkości. Każdy rodzaj nieefektywności wynikający z nieodpowiedniej skali produkcji stanowił do tej pory część nieefektywności technicznej. Na Rysunku 1.4 firma, która produkuje jeden rezultat y z jednego nakładu x znajduje się w punkcie P , gdzie granica produkcji wynosi OT . Dla porównania granica produkcji CRS jest również zaznaczona na rysunku jako prosta OT' .

W modelu zorientowanym na nakłady (gdzie wyniki są stałe, a nakłady ulegają zmianie) efektywność techniczna jest mierzona przy użyciu poziomej odległości danego punktu od granicy produkcji. Efektywność techniczna jest wyrażona jako: $TE_{CRS}^I = y'B'/y'P$. Dla przypadku zorientowania na wyniki (gdzie nakłady są stałe, a wyniki mogą się zmieniać) efektywność techniczna jest wyznaczona jako pionowa odległość od granicy produkcji. W tym wypadku efektywność techniczna jest liczona jako stosunek: $TE_{CRS}^O = x'P/x'A'$. Miary te mogą zostać porównane do tych

wyznaczonych przy zmiennych korzyściach skali (VRS). Dla wariantu orientacji na nakłady efektywność techniczna wynosi teraz: $TE_{VRS}^I = y'B/y'P$, a dla zorientowanego na wyniki: $TE_{VRS}^O = x'P/x'A$.

Rysunek 1.4 Mierzenie efektywności skali produkcji



Dwie rzeczy stają się teraz oczywiste. Po pierwsze, dla wariantu VRS wartość efektywności technicznej zależy, czy została zastosowana orientacja na wyniki, czy na nakłady. Miary te są takie same w przypadku CRS. Po drugie, przy danej orientacji efektywność techniczna firmy jest wyższa przy przyjęciu założenia VRS niż ta obliczona przy założeniu funkcji produkcji o stałych korzyściach skali. Prawidłowość ta jest niezależna od orientacji i dla wszystkich możliwych punktów produkcji znajdujących się wewnątrz lub na granicy funkcji produkcji VRS z wyjątkiem punktu C (gdzie dwie miary będą takie same). Dlatego można stwierdzić, że przy założeniu CRS techniczna nieefektywność wynika z nieodpowiedniego rozmiaru firmy (odcinek $B'B$ dla orientacji na nakłady i AA' dla orientacji na wyniki dla firmy w punkcie P). Dalej można wykazać, że dla firmy w punkcie P :

$$TE_{CRS}^I = \frac{y'B'}{y'P} = \frac{y'B}{y'P} * \frac{y'B'}{y'B},$$

gdzie pierwszy komponent: $\frac{y'B}{y'P} = TE_{VRS}^I$ wyznacza efektywność techniczną zorientowaną na nakłady przy założeniu VRS, a drugi jest miarą efektywności przy założeniu CRS i zorientowaniu na nakłady dla hipotetycznej firmy znajdującej się w punkcie B . Jest to zatem miara nieefektywności wynikająca z rozbieżności między

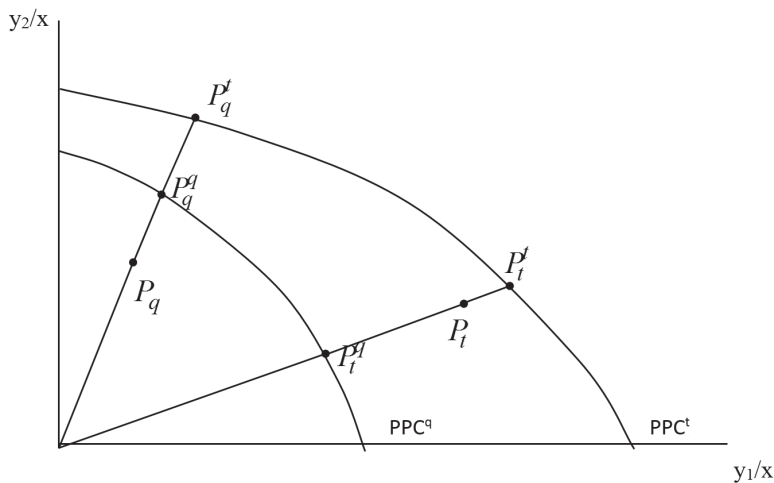
rzeczywistą skalą działalności jednostki w punkcie B i najbardziej produktywną. Może być ona stosowana do oceny efektywności skali produkcji. Dlatego dla orientacji na nakłady całkowita efektywność techniczna firmy w punkcie P jest wyznaczona przez stosunek $y'B'/y'P$, co może być podzielone na efektywność czysto techniczną – $y'B'/y'P$ i efektywność skali produkcji – $y'B'/y'B$. Podobnie przy orientacji na wyniki:

$$TE_{CRS}^O = \frac{x'P}{x'A'} = \frac{x'A}{x'A'} * \frac{x'P}{x'A},$$

całkowita efektywność techniczna firmy jest wyznaczona przez stosunek $x'P/x'A'$, co może być podzielone na efektywność czysto techniczną – $x'P/x'A$ i efektywność skali produkcji – $x'A/x'A'$.

Dotychczas analizowaliśmy efektywność firmy w danym punkcie czasu, dla których różnice w osiąganych wynikach były konsekwencją różnic w efektywnościach. Nie uwzględniano do tej pory roli różnych technologii produkcji. W tym celu rozważmy firmę P , która produkuje dwa wyniki y_1 i y_2 z danego poziomu nakładu x w dwóch okresach czasu, gdzie q to okres początkowy, a t – okres końcowy. Na rysunku 1.5 PPC^q i PPC^t przedstawiają odpowiednio możliwości produkcyjne firmy P w dwóch okresach czasu. W rozważanym okresie nastąpił postęp technologiczny (PPC^t jest przesunięta na zewnątrz w stosunku do PPC^q) w sposób nieneutralny, ponieważ nastąpiło nierównoległe przesunięcie krzywej. Punkt produkcji firmy P przesunął się z Pq do Pt . Ponieważ żaden z tych punktów nie leży na krzywej możliwości produkcyjnych, oznacza to, że firma jest technicznie nieefektywna w obu okresach czasu.

Rysunek 1.5 Pomiar zmian efektywności w czasie, orientacja na wyniki



Rozważmy najpierw oba punkty w czasie oddzielnie. Efektywność techniczna firmy (zorientowana na wyniki według definicji Farella przedstawionej w części 2) w okresie q (TEq) i t (TEt) może zostać zapisana odpowiednio jako: $TE_q = OP_q / OP_q^q$ oraz $TE_t = OP_t / OP_t^t$. Zdefiniujmy funkcję odległości wyników dla okresu q ($D_O^q(x_q, y_{1q}, y_{2q})$), gdzie subskrypt q przy nakładach i wynikach oznacza wielkości z okresu q jako odwrotność maksymalnej wielkości, o którą mogą być zwiększone wyniki (przy stałych nakładach) tak, aby wciąż znajdowały się w zbiorze możliwych do osiągnięcia kombinacji. Jest to wartość mierzona przez TE_q , gdzie $TE_q = D_O^q(x_q, y_{1q}, y_{2q})$. Podobnie dla TEt jest to $TE_t = D_O^t(x_t, y_{1t}, y_{2t})$, gdzie $D_O^t(x_t, y_{1t}, y_{2t})$ oznacza funkcję odległości wyników w okresie t , a indeks t oznacza wielkość nakładów i wyników w okresie t .

Zbadanie, w jaki sposób produktywność firmy zmieniła się w danym okresie czasu, może zostać dokonane na dwa sposoby: dwa punkty produkcji można porównać przy użyciu technologii z okresu q jako technologii bazowej albo technologii z okresu t jako bazowej. Przy pierwszym podejściu efektywność techniczna firmy w punkcie P_t jest mierzona przez porównanie uzyskanych wyników produkcji w okresie t do wyników maksymalnych, jakie mogłyby zostać osiągnięte przy założeniu technologii q (OP_t / OP_t^q , które może być wyrażone jako $D_O^q(x_t, y_{1t}, y_{2t})$) i jest porównane do technicznej efektywności firmy w punkcie, gdzie uzyskany wynik produkcji w okresie q do wyniku, jaki mógłby zostać uzyskany przy założeniu technologii z okresu q (co może zostać zapisane jako: OP_q / OP_q^q i wyrażone: $D_O^q(x_q, y_{1q}, y_{2q})$). Wzrost produktywności w tych dwóch okresach czasu przy użyciu technologii q jako bazowej jest w literaturze nazywany indeksem Malmquista (Malmquist 1963) zorientowanym na wyniki, zdefiniowanym w stosunku do technologii bazowej (M_O^q) i który może zostać zapisany jako:

$$M_O^q = \frac{D_O^q(x_t, y_{1t}, y_{2t})}{D_O^q(x_q, y_{1q}, y_{2q})} = \frac{OP_t / OP_t^q}{OP_q / OP_q^q} . \quad (1.1)$$

Alternatywnie techniczna efektywność firmy w punkcie Pt jest mierzona jako wynik w okresie t w stosunku do maksymalnej produkcji, jaka mogłaby zostać uzyskana przy założeniu technologii z okresu t (może to zostać zapisane jako OP_t / OP_t^t i wyrażone przez odległość: $D_O^t(x_t, y_{1t}, y_{2t})$). Jest on porównywany do efektywności technicznej firmy w punkcie Pq mierzonej jako uzyskane wyniki w okresie q w stosunku do tych, które mogłyby zostać osiągnięte przy technologii z okresu t (OP_q / OP_q^t wyrażone jako: $D_O^t(x_q, y_{1q}, y_{2q})$). Miara wzrostu produktywności pomiędzy dwoma punktami w czasie przy użyciu technologii t jako bazowej jest znana jako indeks Malmquista zorientowany na wyniki, zdefiniowany w stosunku do okresu końcowego (M_O^t) i który może być wyrażony równaniem:

$$M_O^t = \frac{D_O^t(x_t, y_{1t}, y_{2t})}{D_O^t(x_q, y_{1q}, y_{2q})} = \frac{OP_t / OP_t^t}{OP_q / OP_q^t}. \quad (1.2)$$

Dysponujemy dwoma sposobami pomiaru zmian produktywności w czasie, ale nie wiadomo która miara powinna zostać użyta, ponieważ wybór technologii z danego okresu czasu jest arbitralny. Problem ten zostaje rozwiązany poprzez zastosowanie indeksu Malmquista (Mo), który jest średnią geometryczną z indeksów M_O^q i M_O^t (Färe i in. 1994):

$$M_O = [M_O^q \cdot M_O^t]^{\frac{1}{2}} = \left[\frac{D_O^q(x_t, y_{1t}, y_{2t})}{D_O^q(x_q, y_{1q}, y_{2q})} \cdot \frac{D_O^t(x_t, y_{1t}, y_{2t})}{D_O^t(x_q, y_{1q}, y_{2q})} \right]^{\frac{1}{2}} \\ = \left[\frac{OP_t / OP_t^q}{OP_q / OP_q^q} \cdot \frac{OP_t / OP_t^t}{OP_q / OP_q^t} \right]^{\frac{1}{2}}. \quad (1.3)$$

Indeks może zostać przekształcony do postaci:

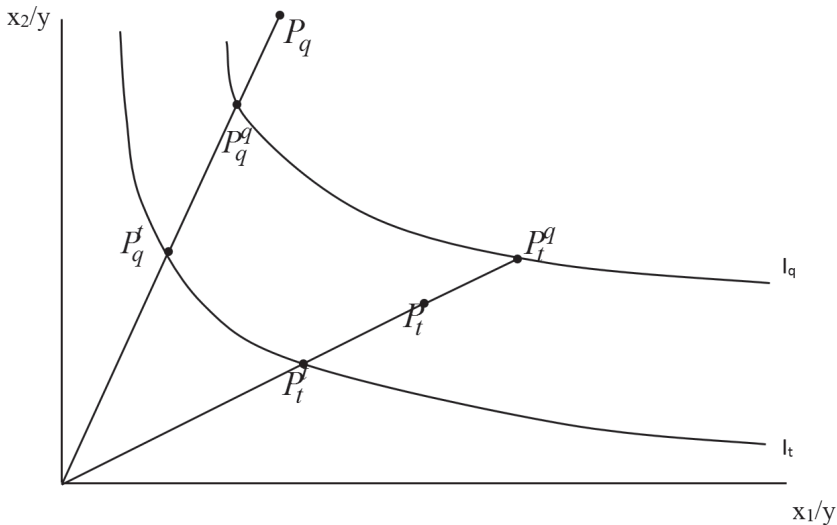
$$M_O = \frac{D_O^t(x_t, y_{1t}, y_{2t})}{D_O^q(x_q, y_{1q}, y_{2q})} \left[\frac{D_O^q(x_t, y_{1t}, y_{2t})}{D_O^t(x_t, y_{1t}, y_{2t})} \cdot \frac{D_O^q(x_q, y_{1q}, y_{2q})}{D_O^t(x_q, y_{1q}, y_{2q})} \right]^{\frac{1}{2}} \\ = \frac{OP_t / OP_t^t}{OP_q / OP_q^q} \left[\frac{OP_t / OP_t^q}{OP_t / OP_t^t} \cdot \frac{OP_q / OP_q^q}{OP_q / OP_q^t} \right]^{\frac{1}{2}}. \quad (1.4)$$

Pierwszym składnikiem równania (1.4) jest stosunek efektywności technicznej w okresie t (okres końcowy) do efektywności technicznej z okresu q (okres początkowy), mierzy on zatem zmianę w efektywności technicznej między tymi dwoma okresami. Jeżeli indeks wynosi 1, oznacza to, że brak jest zmian efektywności technicznej w omawianym czasie. Jeżeli indeks jest większy (mniejszy) od 1, to nastąpił wzrost (spadek) produktywności.

Drugi składnik równania mierzy zmianę w technologii produkcji (tzn. przesunięcie granicy produkcji) między dwoma okresami q i t . Jest średnią geometryczną zmian technologii między dwoma ocenianymi okresami odpowiednio w punktach x_t i x_q . Ten składnik ma wartość 1, gdy nie nastąpiły zmiany w technologii produkcji. Jest większy niż 1 (mniej niż 1), jeśli zmiana technologii produkcji miała pozytywny (negatywny) efekt.

Indeks produktywności Malmquista można również obliczyć dla modelu zorientowanego na nakłady. Musimy wtedy najpierw zdefiniować funkcję odległości zorientowaną na nakłady, która przedstawia, o ile maksymalnie można zmniejszyć nakłady (przy stałych wynikach), pozostając w obrębie możliwego do osiągnięcia zbioru kombinacji produkcji. Na Rysunku 1.6 firma używa teraz dwóch nakładów – x_1 i x_2 do produkcji wyniku, y , I_q i I_t odnoszą się do izokwant odpowiednio z okresu q i t . W omawianym okresie czasu nastąpił postęp technologiczny zobrazowany przez różne położenie: izokwanty I_q i I_t . Produkcja w okresie q i t oznaczona punktami P_q i P_t nie jest efektywna technicznie, ponieważ w obu przypadkach punkty są poza odpowiednimi izokwantami. Odległość funkcji nakładów dla firmy w okresie q wynosi: $D_I^q(x_{1q}, x_{2q}, y_q) = OP_q / OP_q^q$. Jest to odwrotność miary technicznej efektywności Farrell'a orientowanej na nakłady dla firmy P w okresie q . Podobnie funkcja odległości dla firmy P w okresie t jest wyrażona jako: $D_I^t(x_{1t}, x_{2t}, y_t) = OP_t / OP_t^t$.

Rysunek 1.6 Pomiar zmian efektywności w czasie, orientacja na nakłady



Podobnie jak w przypadku podejścia zorientowanego na wyniki pomiar zmian produktywności w czasie może być rozpatrywany dwojako: dwa punkty produkcji mogą być porównywane przy założeniu technologii z okresu q lub z okresu t jako technologii bazowej. W pierwszym przypadku techniczna nieefektywność firmy w punkcie P_q jest obliczana poprzez porównanie nakładu z okresu q w stosunku do minimalnego nakładu przy założeniu technologii z okresu q (tj. $OP_q / OP_q^q = D_I^q(x_{1q}, x_{2q}, y_q)$), a następnie odniesiona do technicznej nieefektywności w punkcie P_t mierzonej jako porównanie

nakładów z okresu t w stosunku do minimalnych nakładów przy założeniu technologii q (tj. OP_t/OP_q^q zapisany jako $D_I^q(x_{1q}, x_{2q}, y_t)$). Pomiar zmian produktywności w czasie przy użyciu technologii z okresu q jako technologii bazowej określony jest indeksem Malmquista (M_I^q), który obliczamy jako:

$$M_I^q = \frac{D_I^q(x_{1q}, x_{2q}, y_q)}{D_I^q(x_{1t}, x_{2t}, y_t)} = \frac{OP_q/OP_q^q}{OP_t/OP_t^q} . \quad (1.5)$$

Natomiast w drugim wariancie nieefektywność firmy w punkcie P_q jest obliczana poprzez porównanie rzeczywiście poniesionych nakładów w okresie q względem wymaganych minimalnych nakładów przy założeniu technologii z okresu t , tj. OP_q/OP_t^t zapisane jako odległość $D_I^t(x_{1q}, x_{2q}, y_q)$. Jest to porównywane z nieefektywnością techniczną firmy w punkcie P_t wyznaczoną przez porównanie nakładów z okresu t do minimalnej wartości nakładów przy zastosowaniu technologii z okresu t (tj. $OP_t/OP_t^t = D_I^t(x_{1t}, x_{2t}, y_t)$). Indeks Malmquista zorientowany na nakłady wyrażony w stosunku do technologii z okresu t (M_I^t) jest dany jako:

$$M_I^t = \frac{D_I^t(x_{1q}, x_{2q}, y_q)}{D_I^t(x_{1t}, x_{2t}, y_t)} = \frac{OP_q/OP_t^t}{OP_t/OP_t^t} . \quad (1.6)$$

Z racji wątpliwości w związku z odpowiedzią na pytanie, z którego okresu powinna być użyta technologia jako bazowa, czy to z okresu początkowego, czy z końcowego, podobnie jak w poprzednim przypadku oblicza się indeks Malmquista, tym razem zorientowany na nakłady M_I jako średnią geometryczną z indeksów M_I^q i M_I^t :

$$M_I = \left[M_I^q \cdot M_I^t \right]^{\frac{1}{2}} = \left[\frac{D_I^q(x_{1q}, x_{2q}, y_q)}{D_I^q(x_{1t}, x_{2t}, y_t)} \cdot \frac{D_I^t(x_{1q}, x_{2q}, y_q)}{D_I^t(x_{1t}, x_{2t}, y_t)} \right]^{\frac{1}{2}} \\ = \left[\frac{OP_q/OP_q^q}{OP_t/OP_t^q} \cdot \frac{OP_q/OP_t^t}{OP_t/OP_t^t} \right]^{\frac{1}{2}} . \quad (1.7)$$

Co może zostać zapisane jako:

$$M_I = \frac{D_I^q(x_{1q}, x_{2q}, y_q)}{D_I^t(x_{1t}, x_{2t}, y_t)} \left[\frac{D_I^t(x_{1t}, x_{2t}, y_t)}{D_I^q(x_{1t}, x_{2t}, y_t)} \cdot \frac{D_I^t(x_{1q}, x_{2q}, y_q)}{D_I^q(x_{1q}, x_{2q}, y_q)} \right]^{\frac{1}{2}}$$

$$= \frac{OP_q / OP_q^q}{OP_t / OP_t^t} \left[\frac{OP_t / OP_t^t}{OP_t / OP_t^q} \cdot \frac{OP_q / OP_q^t}{OP_q / OP_q^q} \right]^{\frac{1}{2}} \quad (1.8)$$

Składowe tego indeksu są interpretowane w sposób odwrotny do komponentów indeksu zorientowanego na wyniki z równania 1.4. W szczególności $D_I^q(x_q, y_q) / D_I^t(x_t, y_t)$ oznacza zmianę technicznej efektywności pomiędzy okresem q i t i wynosi 1. Jeżeli zmiana nie nastąpiła dla wartości mniejszej (większej) od 1 oznacza odpowiednio wzrost (spadek) efektywności technicznej.

Drugi składnik:

$$\left[\frac{D_I^t(x_t, y_t)}{D_I^q(x_t, y_t)} \cdot \frac{D_I^t(x_q, y_q)}{D_I^q(x_q, y_q)} \right]^{\frac{1}{2}},$$

odpowiada zmianom w technologii produkcji pomiędzy okresem q i t , wynosi 1 dla braku zmian, a jest większy niż 1 (mniejszy niż 1), gdy zmiany były pozytywne (negatywne).

W tej części rozdziału przedstawione zostały definicje efektywności, które mogą znaleźć zastosowanie w pomiarze działalności instytucji z sektora edukacji. W praktyce dane potrzebne do analizy funkcji produkcji w kontekście instytucji edukacji dostępne są bądź na poziomie zagregowanym (np. dla szkół, uniwersytetów, na poziomie lokalnych władz odpowiedzialnych za funkcjonowanie sektora edukacji, regionów), bądź na poziomie zdezagregowanym, na poziomie uczniów i studentów. Przedstawione metody mogą być wykorzystywane w każdej z opisanych sytuacji (zarówno przy wykorzystaniu danych zagregowanych i zdezagregowanych), dlatego zostaną one przedstawione oddzielnie w kolejnych dwóch częściach.

3. Techniki pomiaru efektywności: analiza na poziomie jednostki produkcyjnej

Istnieją dwa podstawowe podejścia do pomiaru efektywności: statystyczne (lub ekonometryczne) oraz podejście niestatystyczne (programowanie). Rozróżnienie między tymi dwoma podejściami związane jest z przyjętymi założeniami. Po pierwsze, podejście statystyczne zakłada, że efektywność (różnica pomiędzy osiąganymi przez firmę wynikami a tymi, które mogłyby zostać uzyskane, gdyby firma znajdowała się na granicy produkcji) jest rozłożona według specyficznego rozkładu (Førsund i in. 1980). Podejście statystyczne jest często (ale nie zawsze) podejściem

parametrycznym, dla którego przyjmuje się funkcję produkcji, dla przykładu funkcję produkcji typu Cobba-Douglasa (Sengupta 1999). Parametryczne metody statystyczne za pomocą zapisów matematycznych opisują technologię produkcji i szacują parametry danej funkcji produkcji, których statystyczna istotność wyznaczana jest przy użyciu błędów standardowych (Schmidt 1985/1986). Jednakże w podejściu tym dochodzi do tego, że wszelkie błędy związane np. z nieprawidłową specyfikacją funkcji czy też z błędnymi założeniami co do rozkładu nieefektywności włączone są do miar efektywności (Lovell 1993). Ponadto podejście parametryczne nie jest łatwe do zastosowania w sytuacji, gdy mamy do czynienia z produkcją wielu wyników z wielu nakładów.

W podejściu niestatystycznym nie przyjmuje się założeń na temat rozkładu nieefektywności. Ponadto często (ale nie zawsze) jest ono nieparametryczne, co oznacza, że dane na temat nakładów i wyników są używane do wyznaczenia (poprzez programowanie liniowe) wypukłej powierzchni odpowiadającej granicy efektywności (Sengupta 1999). Podejście niestatystyczne, nieparametryczne pozwala uniknąć problemów związanych z błędną specyfikacją (zarówno funkcji produkcji, jak i rozkładów efektywności), ponieważ w metodach tych rozkłady nie są określane. Dodatkowo metody programowania z łatwością można wykorzystać w sytuacjach, w których występuje zarówno wiele nakładów, jak i wiele wyników. Wadą metod niestatystycznych, nieparametrycznych jest to, że nie dostarczają one oszacowań statystycznej istotności parametrów (Geva-May 2001). Inną wadą jest to, że wypukła powierzchnia jest wyznaczana na podstawie informacji dotyczących ograniczonej liczby obserwacji znajdujących się w próbie.

Wreszcie rozróżnienie między metodami dotyczy ich stochastycznej lub deterministycznej natury (Schmidt 1985/1986; Lovell 1993). Podejście stochastyczne opiera się na założeniu, że odchylenia od funkcji produkcji są wynikiem nie tylko nieefektywności, ale także błędów pomiaru, losowych szoków i zakłóceń statystycznych (Lovell 1993; Ondrich i Ruggiero 2001). Dlatego celem modeli stochastycznych jest podzielenie reszt na dwa składniki: jeden związany z nieefektywnością, drugi będący składnikiem czysto losowym. W praktyce wiąże się to z założeniem określonego rozkładu dla każdego ze składników. Dlatego też metody stochastyczne mają zaletę, że pomiar efektywności nie obejmuje losowych szoków ani błędów pomiarów, ale mogą na nie wpływać błędy złej specyfikacji modelu. Natomiast w metodach deterministycznych przyjmuje się, że każde odchylenie obserwowanego wyniku od granicy produkcji jest wyłącznie konsekwencją nieefektywności (Lovell 1993; Ondrich i Ruggiero 2001). Podczas gdy w metodach deterministycznych brak jest błędów związanych z błędną specyfikacją (ponieważ nie ma w ogóle specyfikacji), to ich wadą jest to, że wszelkie błędy w pomiarze lub błędy stochastyczne są włączone do pomiaru efektywności

Podsumowując, metody pomiaru efektywności mogą być podzielone na statystyczne lub niestatystyczne, parametryczne lub nieparametryczne, deterministyczne lub stochastyczne. Z możliwych wariantów najczęściej występujące to: parametryczne metody statystyczne (deterministyczne lub stochastyczne) i deterministyczne niestatystyczne metody nieparametryczne. W dalszej części tekstu przedstawione zostaną metody najczęściej wykorzystywane wraz z ich rozwojem w kontekście pomiaru efektywności podmiotów edukacji.

Parametryczne metody statystyczne

Odchylenia od granicy są deterministyczne. Załóżmy, że producent i zamienia m nakładów (x) w wynik (y), a proces ten jest opisany równaniem (1.9):

$$y_i = f(x_{i1}, \dots, x_{im}) e^{-u_i} \quad (1.9)$$

dla technologii typu Cobba-Douglasa, można zapisać jako:

$$\ln(y_i) = \ln[f(x_{i1}, \dots, x_{im})] - u_i = \alpha_0 + \sum_{j=1}^m \alpha_j \ln x_{ij} - u_i \quad (1.10)$$

gdzie reszty $-u_i$ spełniają warunek $-u_i \geq 0$ i wyznaczają efektywność producenta i .

Techniczna efektywność producenta i w warunkach zorientowania na wyniki pokazana jest na Rysunku 1.2 i wyznaczona jest jako stosunek aktualnych wyników producenta i do maksymalnych możliwych do osiągnięcia i pokazanych przez granicę produkcji. Dlatego z równania 1.9 efektywność techniczna jest liczona jako:

$$TE_i = \frac{y_i}{f(x_{i1}, \dots, x_{im})} = e^{-u_i} \quad (1.11)$$

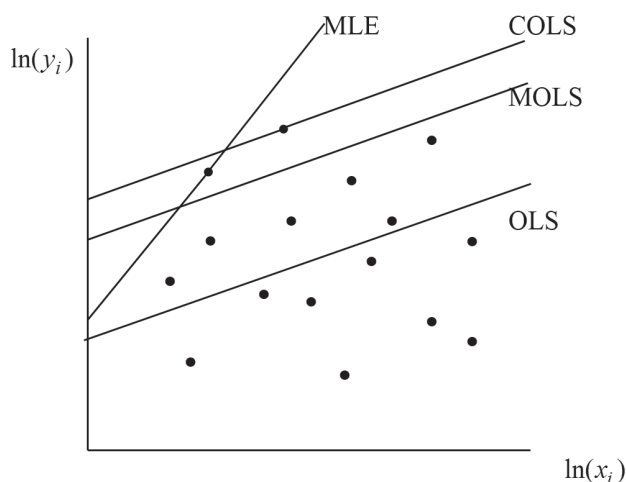
Przy założeniu, że reszty są deterministyczne, istnieją różne możliwe sposoby szacowania granicy produkcji w równaniu (1.9), a tym samym szacowania efektywności technicznej. Po pierwsze, zakładając rozkład dla danej funkcji produkcji (np. Cobba-Douglasa jak w przykładzie powyższym), parametry są szacowane przy użyciu metody najmniejszych kwadratów (MNK). Punkt przecięcia jest następnie przesuwany, do czasu aż wszystkie reszty (oznaczone jako $-u_i$) nie będą dodatnie (tak aby u_i były nieujemne) i co najmniej jedna będzie równa zero³. To podejście jest określane jako metoda najmniejszych kwadratów z korektą (ang. *corrected OLS* – COLS). Podkreślić należy, że w podejściu tym nie określa się rozkładu reszt, a całkowite odstępstwo od granicy produkcji dla danego producenta przypisuje się nieefektywności.

³ W praktyce w pierwszych badaniach na temat edukacji nie dostosowywano wartości reszt, zamiast tego używano estymacji na temat „średniej” funkcji produkcji z niedostosowanymi resztami (np. Johnes 1996; Gray i in. 1984).

W alternatywnych podejściach zakłada się rozkład dla u_i . Najczęściej jest to rozkład półnormalny, chociaż rozkład wykładniczy może być alternatywnie wykorzystany (Lovell, 1993). Parametry równania (1.9) są szacowane przy użyciu MNK i szacowany jest też dodatkowy parametr, a mianowicie średnia z u_i , który jest używany do przesunięcia przecięcia MNK w górę. To podejście jest określane również jako COLS, co jest dość mylące (Førsund i in. 1980) lub jako zmodyfikowana MNK (ang. *modified OLS* – MOLS) (Lovell 1993). Oprócz wad związanych z deterministycznym charakterem reszt podejście MOLS ma taką wadę, że funkcja produkcji może nie być przesunięta wystarczająco daleko, tak żeby wszystkie obserwacje znajdowały się wewnątrz granicy lub na niej. W konsekwencji niektóre reszty mogą mieć odwrotny znak (Førsund i in. 1980; Lovell 1993).

Jako alternatywny estymator można użyć metody największej wiarygodności (ang. *maximum likelihood estimations* – MLE) do oszacowania rozkładu u_i oraz parametrów równania (1.9) (dla przyjętej postaci funkcji produkcji). W konsekwencji funkcja produkcji MLE obejmuje wszystkie obserwacje, ale szacowane parametry różnią się od tych uzyskanych metodą najmniejszych kwadratów (ponieważ związek pomiędzy nakładami i wynikami jest teraz nieliniowy), co pozwala na to, żeby obserwacje efektywne (czyli te, które leżą na granicy produkcji) różniły się pod względem technologii w stosunku do tych, które są wewnątrz granicy (Lovell 1993).

Rysunek 1.7 Porównanie różnych granic produkcji



Wszystkie z tych metod posiadają wady parametrycznych modeli statystycznych. Ponadto należy podkreślić, że COLS i MOLS dają identyczny ranking producentów

co metoda OLS. W końcu, powyżej przedstawione metody są nieodpowiednie do zastosowania w warunkach, gdzie mamy do czynienia z wieloma nakładami i wynikami produkcji⁴. Porównanie tych trzech deterministycznych metod można znaleźć w Lovell (1993) i są one zilustrowane dla przypadku pojedynczego nakładu i pojedynczego wyniku na Rysunku 1.7.

Odległości od granicy są stochastyczne. Dla stochastycznych reszt równanie 1.9 jest zmodyfikowane do następującej formuły:

$$y_i = f(x_{i1}, \dots, x_{im}) e^{\varepsilon_i}, \quad (1.12)$$

gdzie: $\varepsilon_i = v_i - u_i$ a więc równanie (1.10) może zostać zapisane jako:

$$\ln(y_i) = \ln[f(x_{i1}, \dots, x_{im})] + \varepsilon_i = \alpha_0 + \sum_{j=1}^m \alpha_j \ln x_{ij} + \varepsilon_i, \quad (1.13)$$

gdzie $v_i \sim N(0, \sigma_v^2)$, u_i i v_i są statystycznie niezależne, a $u_i \geq 0$ (Aigner in. 1977). Reszty są podzielone na dwa składniki. Jeden jest związany z błędem pomiaru i zmianami losowymi (v_i), podczas, gdy drugi składnik, najczęściej o rozkładzie wykładniczym lub pół normalnym, odpowiada technicznej nieefektywności (u_i). Parametry funkcji produkcji mogą być oszacowane metodą MOLS (Førsund i in. 1980; Lovell 1993) lub MLE. Technika ta zwana jest stochastyczną metodą graniczną. Jeżeli (u_i) ma rozkład półnormalny, to przykładowo logarytm funkcji największej wiarygodności (ang. *log-likelihood function*) jest zapisany jako:

$$\ln L = -n \ln \sigma - \frac{n}{2} \ln \frac{2}{\pi} - \frac{1}{2} \sum_i \frac{\varepsilon_i^2}{2\sigma^2} + \sum_i \ln \Phi\left(\frac{-\varepsilon_i \lambda}{\sigma}\right), \quad (1.14)$$

gdzie: $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \sigma_u / \sigma_v$ i Φ jest dystrybuantą standardowego rozkładu normalnego.

Możliwe jest obliczenie błędów ε_i jako odchyień między wynikami obserwowanymi a wynikami estymowanymi. Dla warunku: $u_i \sim N(0, \sigma_u^2)$ średnia techniczna nieefektywność wynosi $\sigma_u \sqrt{(2/\pi)}$ i może zostać wyznaczona na podstawie oszacowanego σ_u . Metoda stochastycznej granicy pozwala uzyskać informację na temat technicznej efektywności każdej jednostki za pomocą metody zapoczątkowanej przez Jondrow i in. (1982). Dla normalnego i półnormalnego rozkładu reszt szacunki dla jednostek są otrzymywane w następujący sposób⁵.

⁴ Zobacz np. Lovell (1993), jak zaadaptować te metody do przypadku wielu nakładów i wielu wyników.

⁵ Jondrow i in. (1982) podają także rozwiązanie dla rozkładu normalno-wykładniczego.

Zauważmy, że ε_i jest obserwowane $\forall i$, a więc zadanie polega na podziale oszacowanego ε_i , dla każdego i na składnik u_i oraz v_i ⁶. Można wykazać, że warunkowy rozkład u_i dla ε_i to $N(\mu_{*i}, \sigma_{*i}^2)$ odcięty w punkcie 0, gdzie $\mu_{*i} = -\varepsilon_i \sigma_u^2 / \sigma^2$ i $\sigma_{*i}^2 = \sigma_u^2 \sigma_v^2 / \sigma^2$. Punktowy estymator u_i jest dany przez średnią rozkładu normalnego:

$$E(u_i / \varepsilon_i) = \frac{\sigma_u^2 \sigma_v^2}{\sigma^2} \left[\frac{\phi(\varepsilon_i \lambda / \sigma)}{1 - \Phi(\varepsilon_i \lambda / \sigma)} - \frac{\varepsilon_i \lambda}{\sigma} \right], \quad (1.15)$$

gdzie Φ jest standardowym rozkładem normalnym. Ponieważ μ_{*i} i σ_{*i}^2 są nieznane, zastosowanie tej metody wymaga użycia ich wartości estymowanych, błąd wynikający z tego przybliżenia zanika asymptotycznie i dlatego dla wystarczająco dużych prób można go nie brać pod uwagę.

Analiza granicy stochastycznej jest atrakcyjna ze względu na jej charakter statystyczny i ugruntowanie w teorii ekonomii. Jej krytyka wiąże się z dość szerokimi przedziałami ufności odnośnie oszacowanych wskaźników nieefektywności. Przedziały te mogą być zbyt szerokie, aby metoda ta zyskała przychylność praktyków, analitycy mogą zadać także pytania na temat poprawności rozróżnienia pomiędzy dwoma składnikami reszt. Oczywiście jest to kwestia do rozstrzygnięcia w badaniach empirycznych. Dodatkowo przyjęcie danej funkcji rozkładu dla reszt (czy to półnormalnego, czy wykładniczego) w przypadku składnika określającego nieefektywność techniczną jest założeniem, które nie ma swojego umocowania w teorii. Nie ma więc powodu, aby preferować jeden rodzaj rozkładu nad drugi, a błędy specyfikacji stają się częścią miary efektywności. W przypadku, w którym efektywność organizacji jest wynikiem działalności wielu osób o różnej efektywności, można się spodziewać, że rozkład efektywności instytucji będzie rozkładem normalnym. Jednak analiza granicy stochastycznej nie może być używana do oceny efektywności, jeżeli rozkład u_i jest normalny, ponieważ w takim przypadku nie będzie go można odróżnić od v_i .

Deterministyczne niestatystyczne metody nieparametryczne

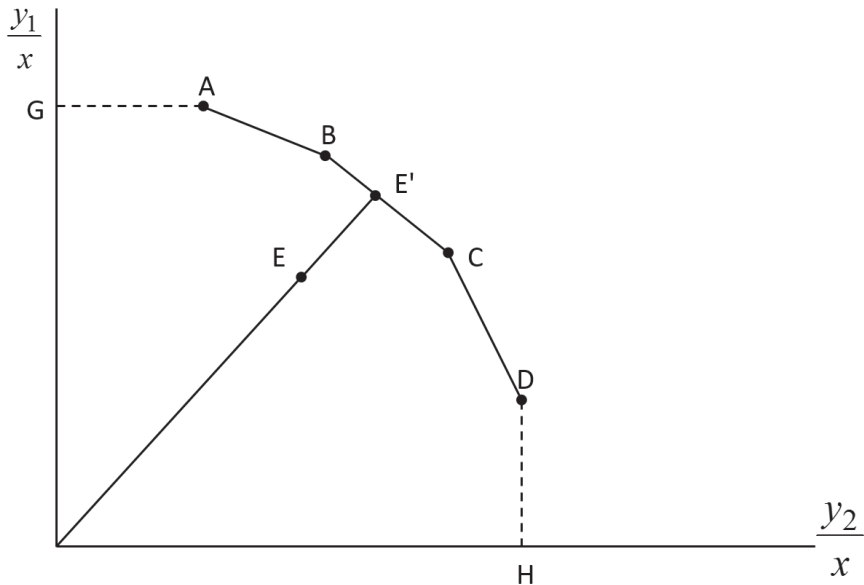
W części tej zostaną opisane deterministyczne, niestatystyczne metody nieparametryczne. Wśród nich metoda najczęściej stosowana do pomiaru efektywności technicznej, czyli analiza obwiedni danych (ang. *data envelopment analysis*, DEA). Podstawowe założenia metody DEA przedstawiono poniżej.

DEA ze stałymi korzyściami skali. Celem metody DEA opracowanej przez Charnes i in. (1978) w oparciu o prace Dantzig (1951) i Farrell (1957) jest

⁶ Pierwszy ze składników jest miarą nieefektywności technicznej (oznaczonej przez u_i w równaniu (1.3)) i może posłużyć do znalezienia technicznej efektywności dla każdego producenta poprzez wprowadzenie do równania (1.12).

oszacowanie granicy możliwości produkcyjnych, a tym samym dokonanie oceny technicznej efektywności każdej firmy lub jednostki decyzyjnej (ang. *decision making unit*, DMU) w stosunku do granicy możliwości produkcyjnych. W najprostszej wersji DEA zakłada stałe korzyści skali (CRS).

Rysunek 1.8 Ilustracja metody DEA



Rozważmy najpierw prosty przykład pięciu uniwersytetów (A, B, C, D, E) produkujących dwa rezultaty – y_1 (liczba absolwentów osiągających wystarczająco „dobre” stopnie) i y_2 (liczba absolwentów wchodzących na rynek pracy) z jednego nakładu x (liczba studentów). Stosunek wyników y_1 do nakładu można zilustrować względem ilorazu: wynik y_2 do nakładu jak na Rysunku 1.8. Prosta, która łączy uniwersytety A, B, C i D, jest granicą efektywności, ponieważ żadna z jednostek znajdujących się na granicy nie jest w stanie wytworzyć więcej rezultatów (przy danym nakładzie), a więc wszystkie uniwersytety znajdujące się na granicy są efektywne technicznie. Jednak uniwersytet E jest położony wewnątrz granicy. Promień wyprowadzony ze środka układu współrzędnych przez punkt E i przeciągnięty do granicy, którą przecina w punkcie E', wskazuje, że DMU z połączonych jednostek B i C mógłby wyprodukować więcej obu wyników (przy danym nakładzie) niż jednostka E. W rzeczywistości wyniki osiągnięte przez uniwersytet E mogą być proporcjonalnie zwiększone (bez zwiększania nakładu) o wartość: $1 - OE/OE'$, gdzie OE/OE' ukazuje względny

poziom efektywności jednostki E w stosunku do innych jednostek w próbie, a miara zgodna jest z definicją efektywności technicznej zorientowanej na wyniki Farella. W alternatywnym podejściu, gdzie wyznaczono stosunek nakładów do wyników, granica efektywności pokazana jako SS' na Rysunku 1.1, można wyliczyć efektywność techniczną zorientowaną na nakłady. Dla CRS miary efektywności technicznej zorientowane na nakłady i wyniki są tożsame.

W praktyce DMU mogą wytwarzać wiele wyników z wielu nakładów, dlatego wykorzystuje się techniki programowania w celu wyznaczenia granicy, która łączy wszystkie efektywne DMU, granica SS' na Rysunku 1.1 (jeżeli zastosowano orientację na nakłady) lub DD' z Rysunku 1.2 (orientacja na wyniki). Techniczna efektywność jednostki k jest zdefiniowana jako stosunek ważonej sumy nakładów do ważonej sumy wyników (Charnes i in. 1978):

$$TE_k = \frac{\sum_{r=1}^s u_r y_{rk}}{\sum_{i=1}^m v_i x_{ik}}, \quad (1.16)$$

gdzie s to wyniki, a m nakłady; y_{rk} to r -ty wynik osiągnięty przez k -tą jednostkę decyzyjną; x_{ik} to i -ty nakład wykorzystany przez k -tą jednostkę decyzyjną, u_r oznacza wagę dla r -tego wyniku; a v_i to waga użyta w stosunku do i -tego nakładu. TE_k odpowiada efektywności technicznej zdefiniowanej w części 2, ale należy zauważyć, że jest to miara relatywna, liczona w stosunku do pozostałych jednostek decyzyjnych z próby.

Jednostka decyzyjna k maksymalizuje swój wskaźnik efektywności z zastrzeżeniem pewnych ograniczeń: (i) wagi są uniwersalne: co oznacza, że waga użyta w stosunku do jednostki k , kiedy jest zastosowana do każdego innego DMU z próby, to uzyskany na ich podstawie wskaźnik efektywności nie może być większy od 1; (ii) wagi dla nakładów i dla rezultatów są dodatnie⁷. Dla każdej n -tej jednostki decyzyjnej następujący problem programowania liniowego musi zostać rozwiązany:

$$\text{maksymalizacja: } \frac{\sum_{r=1}^s u_r y_{rk}}{\sum_{i=1}^m v_i x_{ik}} \quad (1.17)$$

⁷ Warunek, aby wagi były dodatnie (w odróżnieniu od nieujemnych), został wprowadzony przez Charnesa i in. (1979), gdzie badacze ci rozwinęli wcześniejszy model (Charnes i in. 1978), w którym wymagali żeby wagi były nieujemne. Wymóg dodatniości wag gwarantuje, że efektywne DMU (ze wskaźnikiem efektywności równym 1) nie ma luzów ani na nakładach, ani na wynikach.

$$\text{przy ograniczeniach:} \quad \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1 \quad j = 1, \dots, n \quad (1.18)$$

$$u_r, v_i > 0 \quad \forall r = 1, \dots, s; i = 1, \dots, m$$

Ten problem programowania liniowego może zostać przekształcony w ramach maksymalizacji ważonej sumy wyników przy stałych wartościach nakładów (podejście zorientowane na wyniki) lub przez minimalizację ważonej sumy nakładów przy stałych wynikach (podejście zorientowane na nakłady). Zadanie pierwotne⁸ dla każdego z podejść podano poniżej:

**Orientacja na wyniki (CRS)
zadanie pierwotne:**

$$\text{Minimalizacja} \quad \sum_{i=1}^m v_i x_{ik} \quad (1.19a)$$

przy warunkach (1.20a)

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} \geq 0 \quad j = 1, \dots, n$$

$$\sum_{r=1}^s u_r y_{rk} = 1$$

$$u_r, v_i > 0 \quad \forall r = 1, \dots, s; i = 1, \dots, m$$

**Orientacja na nakłady (CRS)
zadanie pierwotne:**

$$\text{Maksymalizacja} \quad \sum_{r=1}^s u_r y_{rk} \quad (1.19b)$$

przy warunkach (1.20b)

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} \geq 0 \quad j = 1, \dots, n$$

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

$$u_r, v_i > 0 \quad \forall r = 1, \dots, s; i = 1, \dots, m$$

W praktyce często wykorzystuje się zadanie dualne (równania dla każdego z podejść są podane poniżej), które jest prostsze do obliczenia niż równanie pierwotne, posiada $s+m$ warunków ograniczających w stosunku do $n+1$ w równaniu pierwotnym.

⁸ Terminologia zadania pierwotnego odpowiada pracy Charnesa i in. (1978). Należy jednak zauważyć, że maksymalizacja rezultatów przy danych nakładach (tj. zorientowanie na wyniki) jest osiągnięta poprzez pierwotne zadanie minimalizacji funkcji celu (Norman i Stoker 1991).

**Orientacja na wyniki
zadanie dualne (CRS):**

Maksymalizacja ϕ_k (1.21a)
przy warunkach (1.11a)

$$\phi_k y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} \leq 0 \quad r = 1, \dots, s$$

$$x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, m$$

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

**Orientacja na nakłady
zadanie dualne(CRS):**

Minimalizacja θ_k (1.21b)
przy warunkach (1.22b)

$$y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} \leq 0 \quad r = 1, \dots, s$$

$$\theta_k x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, m$$

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

W poprzednich zapisach nie braliśmy pod uwagę luzów w pomiarach efektywności. Problem luzów powstaje, ponieważ część granicy efektywności przebiega równoległe do osi pionowej i poziomej (odcinki *GA* i *DH* na Rysunku 1.8). Załóżmy, że dodano dodatkową jednostkę decyzyjną: *F*, która na Rysunku 1.8 znajduje się na odcinku *DH*. Punkt ten znajduje się na granicy, a więc jest efektywny z wynikiem równym 1. Jednakże w próbie istnieje np. jednostka *D*, która wytwarza ten sam poziom wyników y_2 , ale więcej wyników y_1 w stosunku do tych samych nakładów. Jednostka decyzyjna *F* może zatem zwiększyć swoją efektywność pod względem jednego z produktów, a więc mówi się, że posiada luzy rezultatów. Dla orientacji na nakłady luz będzie oznaczał, że jest możliwym zmniejszenie chociaż jednego nakładu bez zmian osiągniętych wyników. Powyższe równanie może zostać zapisane, biorąc pod uwagę luzy nakładów i wyników (odpowiednio oznaczone jako s_r i s_i):

**Orientacja na wyniki
zadanie dualne z luzami (CRS):**

Maksymalizacja $\phi_k - \varepsilon \sum_{r=1}^s s_r - \varepsilon \sum_{i=1}^m s_i$
(1.23a)
przy warunkach (1.24a)

$$\phi_k y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} + s_r = 0 \quad r = 1, \dots, s$$

$$x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} - s_i = 0 \quad i = 1, \dots, m$$

$$\lambda_j, s_r, s_i \geq 0$$

$$\forall j = 1, \dots, n; r = 1, \dots, s; i = 1, \dots, m$$

**Orientacja na nakłady
zadanie dualne z luzami (CRS):**

Minimalizacja $\theta_k - \varepsilon \sum_{r=1}^s s_r - \varepsilon \sum_{i=1}^m s_i$
(1.23b)
przy warunkach (1.24b)

$$y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} + s_r = 0 \quad r = 1, \dots, s$$

$$\theta_k x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} - s_i = 0 \quad i = 1, \dots, m$$

$$\lambda_j, s_r, s_i \geq 0$$

$$\forall j = 1, \dots, n; r = 1, \dots, s; i = 1, \dots, m$$

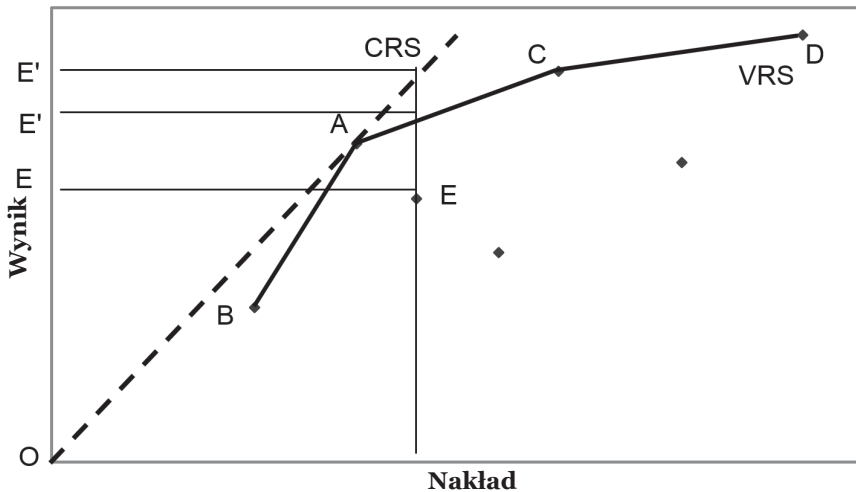
Jednostka decyzyjna k jest efektywna, jeżeli wskaźnik efektywności $TE_k = \frac{1}{\phi_k} = 1$ (równocześnie $TE_k = \theta_k = 1$), a luzy: $s_r, s_i = 0$.

DEA ze zmiennymi korzyściami skali. Założenie CRS może zostać zdjęte, a model DEA można łatwo zmodyfikować, aby uwzględnić zmienne korzyści skali (VRS) (Banker i in. 1984). Założenie VRS w przeciwieństwie do CRS wpływa na wartości wskaźników efektywności DMU, jak pokazano na Rysunku 1.9. Granica efektywności CRS (linia przerywana) i granica efektywności VRS (linia ciągła $BACD$) odpowiada sytuacji, gdzie jest tylko jeden nakład i jeden wynik. Rozważmy jednostkę E , której wskaźnik efektywności jest równy stosunkowi OE/OE' przy VRS i jest wyższy od wskaźnika dla CRS wynoszącego: OE/OE'' ⁹.

Należy zauważyć, że zbiór DMU zidentyfikowanych jako nieefektywne w ramach VRS będzie taki sam bez względu na orientację modelu: czy to zorientowany na nakłady, czy na wyniki. Jednak wartość wskaźnika efektywności dla nieefektywnego DMU będzie się różniła w zależności od wyboru orientacji modelu (orientacja na nakłady czy na wyniki), co pokazano na Rysunku 1.4. Jest to odmienne od sytuacji CRS, gdzie wybór orientacji nie wpływa na wynik efektywności nieefektywnego DMU.

W warunkach VRS dla każdego DMU należy rozwiązać poniższe równania programowania liniowego:

Rysunek 1.9 Granice DEA: stałe korzyści skali versus zmienne korzyści skali



⁹ Wskaźniki efektywności są otrzymywane na podstawie pionowej odległości od granicy i dlatego oznaczają orientację na wyniki. Wskaźniki efektywności mogłyby zostać wyznaczone na podstawie odległości poziomej od granicy i oznaczałyby zorientowanie na nakłady.

**Orientacja na wyniki (VRS)
zadanie pierwotne:**

$$\text{Minimalizacja } \sum_{i=1}^m v_i x_{ik} - c_k \quad (1.25a)$$

przy warunkach (1.26a)

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} \geq 0 \quad j = 1, \dots, n$$

$$\sum_{r=1}^s u_r y_{rk} = 1$$

$$u_r, v_i > 0 \quad \forall r = 1, \dots, s; i = 1, \dots, m$$

**Orientacja na nakłady (VRS)
zadanie pierwotne:**

$$\text{Maksymalizacja } \sum_{r=1}^s u_r y_{rk} \quad (1.25b)$$

przy warunkach (1.26b)

$$\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s u_r y_{rj} \geq 0 \quad j = 1, \dots, n$$

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

$$u_r, v_i > 0 \quad \forall r = 1, \dots, s; i = 1, \dots, m$$

gdzie c_k jest miarą zwrotu ze skali produkcji dla jednostki k , a dualne zadanie przedstawiono poniżej:

**Orientacja na wyniki (VRS)
zadanie dualne:**

$$\text{Maksymalizacja } \phi_k \quad (1.27a)$$

przy warunkach (1.28a)

$$\phi_k y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} \leq 0 \quad r = 1, \dots, s$$

$$x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, m$$

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

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

**Orientacja na nakłady (VRS)
zadanie dualne:**

$$\text{Minimalizacja } \theta_k \quad (1.27b)$$

przy warunkach (1.28b)

$$y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} \leq 0 \quad r = 1, \dots, s$$

$$\theta_k x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, m$$

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

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

Po dodaniu luzów, otrzymujemy:

**Orientacja na wyniki (VRS)
zadanie dualne:**

$$\text{Maksymalizacja: } \phi_k - \varepsilon \sum_{r=1}^s s_r - \varepsilon \sum_{i=1}^m s_i \quad (1.29a)$$

przy warunkach (1.30a)

$$\phi_k y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} + s_r = 0 \quad r = 1, \dots, s$$

$$x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} - s_i = 0 \quad i = 1, \dots, m$$

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

$$\lambda_j, s_r, s_i \geq 0$$

$$\forall j = 1, \dots, n; r = 1, \dots, s; i = 1, \dots, m$$

**Orientacja na nakłady (VRS)
zadanie dualne:**

$$\text{Minimalizacja } \theta_k - \varepsilon \sum_{r=1}^s s_r - \varepsilon \sum_{i=1}^m s_i \quad (1.29b)$$

Przy warunkach (1.30b)

$$y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} + s_r = 0 \quad r = 1, \dots, s$$

$$\theta_k x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} - s_i = 0 \quad i = 1, \dots, m$$

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

$$\lambda_j, s_r, s_i \geq 0$$

$$\forall j = 1, \dots, n; r = 1, \dots, s; i = 1, \dots, m$$

Całkowita efektywność techniczna jednostki k jest mierzona jako $TE_k = 1/\phi_k$ (dla orientacji na wyniki) oraz $TE_k = \theta_k$ (dla orientacji na nakłady), a efektywność skali może być obliczona dla jednostki k jako:

$$SCE_k = \frac{TE_{k,CRS}}{TE_{k,VRS}} \quad (1.31)$$

Ponadto można określić czy dana jednostka działa w obrębie rosnących lub malejących korzyści skali poprzez oszacowanie, oprócz modeli CRS i VRS, modelu z nierosnącymi korzyściami skali (ang. *non-increasing returns to scale* – NIRS)¹⁰ i porównanie ich rezultatów (Coelli i in. 1998).

DEA i efektywność alokacyjna. Powyżej przedstawione założenia metody DEA skupiały się na pomiarze efektywności technicznej. Jednak metoda DEA może dostarczyć informacji o całkowitej efektywności, efektywności alokacyjnej dla każdego DMU zgodnie z definicją z części 2. W celu wyznaczenia efektywności alokacyjnej niezbędna jest wiedza na temat cen wszystkich nakładów (w_j) i wszystkich wyników (p_j). W przypadku szkół i instytucji szkolnictwa wyższego takie informacje są rzadko

¹⁰ Jest to VRS, z tym że warunek równości $\sum_{j=1}^n \lambda_j = 1$ jest zastąpiony przez nierówność $\sum_{j=1}^n \lambda_j \leq 1$.

dostępne, a zatem metodologia określania efektywności alokacyjnej zostanie przedstawiona w sposób bardzo zwięzły.

Efektywność alokacyjna związana jest z efektywnością całkowitą i efektywnością techniczną w następujący sposób:

$$AE_k = \frac{OE_k}{TE_k} \quad (1.32)$$

i może zostać wyznaczona albo przez orientację na nakłady (minimalizacja kosztów), albo przez orientację na wyniki (maksymalizacja przychodów) – zobacz część 2. Dla maksymalizacji przychodów pierwszym etapem jest obliczenie technicznej efektywności (TE_k) przy użyciu orientacji na wyniki (CRS albo VRS). W kolejnym kroku rozwiązuje się maksymalizację przychodów DEA. Poniżej przedstawiono równania dla wariantu VRS (równania od (1.33a) do (1.34a)):

Orientacja na wyniki (VRS):

Maksymalizacja: $\sum_{r=1}^s p_{rk} y_{rk}^*$ (1.33a)

przy warunkach: (1.34b)

$$y_{rk}^* - \sum_{j=1}^n \lambda_j y_{rj} \leq 0 \quad r = 1, \dots, s$$

$$x_{ik} - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, m$$

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

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

gdzie p_{rk} jest ceną rezultatu r -tego ($r = 1, \dots, s$) dla jednostki k , y_{rk}^* jest obliczone na podstawie programowania liniowego i jest wielkością maksymalizującą przychód dla rezultatu r ($r = 1, \dots, s$) dla DMU k , gdzie nakłady wynoszą x_{ik} , a ceny rezultatów: p_{rk} .

Orientacja na nakłady (VRS):

Maksymalizacja: $\sum_{i=1}^m w_{ik} x_{ik}^*$ (1.33b)

przy warunkach: (1.34b)

$$y_{rk} - \sum_{j=1}^n \lambda_j y_{rj} \leq 0 \quad r = 1, \dots, s$$

$$x_{ik}^* - \sum_{j=1}^n \lambda_j x_{ij} \geq 0 \quad i = 1, \dots, m$$

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

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

gdzie w_{ik} jest ceną nakładu i -tego ($i = 1, \dots, m$) dla jednostki k , x_{rk}^* jest obliczone na podstawie programowania liniowego i jest wielkością minimalizującą koszty dla nakładu i ($i = 1, \dots, m$) i jednostki k , gdzie wyniki y_{rk} , a ceny nakładów: w_{ik} .

Całkowita efektywność jednostki k , tak jak zdefiniowano ją w części 2, może zostać obliczona jako:

$$OE_k = \frac{\sum_{r=1}^s p_{rk} y_{rk}}{\sum_{r=1}^s p_{rk} y_{rk}^*} \quad (1.35)$$

Efektywność alokacyjną można obliczyć za pomocą równania (1.32).

W przypadku minimalizacji kosztów pierwszym etapem jest obliczenie technicznej efektywności (TE) przy użyciu orientacji na nakłady (CRS albo VRS). W kolejnym kroku rozwiązuje się minimalizację kosztów DEA. Równania dla wariantu VRS przedstawiono powyżej (równania od (1.33b) do (1.34b)). Całkowita efektywność dla DMU k , jak zdefiniowano ją w części 2, jest obliczana jako:

$$OE_k = \frac{\sum_{i=1}^m w_i k x_{ik}}{\sum_{i=1}^m w_i k x_{ik}^*} \quad (1.36)$$

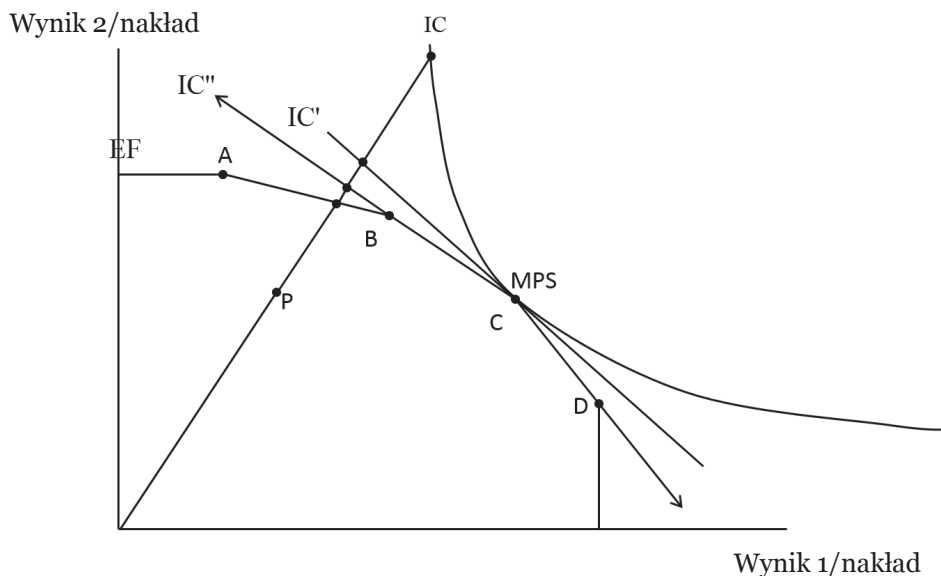
Efektywność alokacyjna może być znów wyznaczona za pomocą równania (1.32).

DEA i preferencje. Podejmowano różne próby w celu włączenia informacji dotyczących preferencji do metody DEA. Najprostszą metodą jest nałożenie ograniczeń na wagi przypisane poszczególnym nakładom i wynikom. Na przykład przy ustawianiu celów na rezultaty (dla danych nakładów) można określić, żeby jeden z wyników miał znaczenie priorytetowe. W efekcie oznacza to utrzymywanie nakładów i niektórych (ale nie wszystkich) wyników na stałym poziomie w procedurze ustawiania celu (zob. np. Thanassoulis i Dunstan 1994). Taka procedura idzie w pewnym sensie w kierunku parametryzacji funkcja produkcji.

Ostatnio Halme i in. (1999) dostosowali DEA w celu włączenia preferencji „decydenta” (ang. *decision maker* DM). Procedura jest zilustrowana dla wariantu orientacji na wyniki na Rysunku 1.10, gdzie pięć jednostek produkuje dwa wyniki z pojedynczego nakładu. Pierwszym etapem jest standardowa DEA w celu identyfikacji granicy (ABCD na rysunku 1.10). A zatem jednostka w punkcie P ma efektywność techniczną mierzoną przez stosunek $TE=OP/OP_1$. Następnym krokiem jest zidentyfikowanie dla jednostki najkorzystniejszej kombinacji nakładów i wyników zwanych najbardziej preferowanym rozwiązaniem (ang. *most preferred solution* – MPS), które leży na granicy efektywności. Halme i in. (1999) użyli procedury szukania w formie

wieloobiektowego programowania liniowego w celu lokalizacji MPS, ale możliwe jest użycie alternatywnych procedur.

Rysunek 1.10 Pomiar efektywności wartości



Ostatnim krokiem jest obliczenie miary efektywności, która obejmuje preferencje jednostki nazwanej przez Halme i in. (1999) „efektywnością wartości” (ang. *value efficiency* – VE). Zakłada się, że MPS leży również na (nieznanej) krzywej obojętności odzwierciedlającej wszystkie punkty jednakowej użyteczności (oznaczone przez IC na Rysunku 1.10). Prawdziwa wartość efektywności, która zawiera preferencje jednostki, wyrażona jest stosunkiem: $VE = OP/OP_4$, ale ponieważ krzywa obojętności nie jest obserwowana, nie można wyznaczyć tego ilorazu. Zakłada się, że wszystkie punkty na krzywej obojętności znajdują się powyżej linii, która jest styczną do granicy efektywności w MPS (linia styczna jest oznaczona jako IC' na Rysunku 1.10). Styczna ta może zostać wykorzystana do oszacowania efektywności wartości – OP/OP_3 (Korhonen i in. 2001). W praktyce nie można oszacować efektywności wartości, zamiast tego należy użyć prostej IC'' na Rysunku 1.10, aby oszacować przybliżenie efektywności wartości jako – $VE^* = OP/OP_2$, która daje wartość zawyżoną w górę¹¹. Oszacowanie VE^* dla wszystkich DMU wymaga znalezienia rozwiązania równań programowania liniowego podobnych do tych stosowanych w standardowej procedurze DEA (zob. Korhonen i in. 2001).

¹¹ Alternatywnie prosta IC''' wprowadzona z DC przecina promień OP w punkcie P_4 i $VE^{**} = OP/OP_4$ daje bardziej konserwatywne (ale tak samo poprawne) miary.

DEA, efektywność techniczna i zmiany technologii produkcji. W części 2 opisano, w jaki sposób indeks produktywności Malmquista (zorientowany na nakłady lub na wyniki) może być wykorzystany do wyznaczenia zmian efektywności w czasie w podziale na zmiany efektywności technicznej i na zmiany w technologii produkcji. DEA może być łatwo dostosowana do obliczania wymaganego indeksu produktywności Malmquista. Dla przykładu bliższe przyjrzenie się indeksowi Malmquista zorientowanemu na wyniki¹² (zob. równanie 1.37) wskazuje, że trzeba obliczyć cztery funkcje odległości:

$$M_O(x_t, y_t, x_q, y_q) = \frac{D_O^t(x_t, y_t)}{D_O^q(x_q, y_q)} \left[\frac{D_O^q(x_t, y_t)}{D_O^t(x_t, y_t)} \cdot \frac{D_O^q(x_q, y_q)}{D_O^t(x_q, y_q)} \right]^{\frac{1}{2}}, \quad (1.37)$$

gdzie: x_q i x_t to wektory nakładów użyte odpowiednio w okresie q i t , a y_q i y_t to wektory wyników z okresu q i t . $D_O^q(x_q, y_q)$ i $D_O^t(x_t, y_t)$ są znajdowane przez rozwiązanie DEA dla wariantu zorientowania na wyniki CRS oddzielnie dla poszczególnych punktów w czasie. W szczególności:

Obliczanie $D_O^t(x_t, y_t)$:

$$\frac{1}{D_O^t(x_t, y_t)} = \text{Maksymalizacja } \phi_k \quad (1.38a)$$

przy warunkach (1.39a)

$$\phi_k y_{rkt} - \sum_{j=1}^n \lambda_j y_{rjt} \leq 0 \quad r = 1, \dots, s$$

$$x_{ikt} - \sum_{j=1}^n \lambda_j x_{ijt} \geq 0 \quad i = 1, \dots, m$$

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

Obliczanie $D_O^q(x_q, y_q)$:

$$\frac{1}{D_O^q(x_q, y_q)} = \text{Maksymalizacja } \phi_k \quad (1.38b)$$

przy warunkach (1.39b)

$$\phi_k y_{rkq} - \sum_{j=1}^n \lambda_j y_{rjq} \leq 0 \quad r = 1, \dots, s$$

$$x_{ikq} - \sum_{j=1}^n \lambda_j x_{ijq} \geq 0 \quad i = 1, \dots, m$$

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

¹² W tekście przedstawiona została tylko metoda do obliczenia indeksu Malmquista zorientowanego na wyniki. DEA zorientowana na nakłady może zostać użyta w celu obliczenia indeksu Malmquista zorientowanego na nakłady.

gdzie subskrypty q i t dotyczą odpowiednio okresu q i t . Ostatnie dwie odległości przy orientacji na wyniki zostają znalezione poprzez rozwiązanie następujących równań programowania liniowego:

Obliczanie $D_O^t(x_q, y_q)$:

$$\frac{1}{D_O^t(x_q, y_q)} = \text{Max } \phi_k \quad (1.40a)$$

przy warunkach (1.41a)

$$\phi_k y_{rkq} - \sum_{j=1}^n \lambda_j y_{rjt} \leq 0 \quad r = 1, \dots, s$$

$$x_{ikq} - \sum_{j=1}^n \lambda_j x_{ijt} \geq 0 \quad i = 1, \dots, m$$

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

Obliczanie $D_O^q(x_t, y_t)$:

$$\frac{1}{D_O^q(x_t, y_t)} = \text{Max } \phi_k \quad (1.40b)$$

przy warunkach (1.41b)

$$\phi_k y_{rkt} - \sum_{j=1}^n \lambda_j y_{rjq} \leq 0 \quad r = 1, \dots, s$$

$$x_{ikt} - \sum_{j=1}^n \lambda_j x_{ijq} \geq 0 \quad i = 1, \dots, m$$

$$\lambda_r \geq 0 \quad \forall r = 1, \dots, s$$

Należy zauważyć, że ϕ s i λ s nie muszą mieć takich samych wartości w przedstawionych zestawach równań. Ponadto wartość optymalizacji wynikająca z dwóch ostatecznych LP nie może być większa lub równa 1, podobnie jak w standardowej procedurze DEA nakierowanej na wyniki. W przypadku tych dwóch programowań liniowych punkt produkcji z danego okresu czasu jest porównywany z granicą produkcji z innego okresu (mającą inną technologię), a zatem może znajdować się poza osiągalnym zestawem produkcji w danym okresie. Jest to szczególnie prawdopodobne, gdy punkt produkcji dla końcowego okresu czasu (w tym przypadku t) jest porównywany z granicą z początkowego okres (w tym przypadku q).

Bootstrapping. Jednym z głównych zastrzeżeń dotyczących metody DEA jest jej niezdolność do zapewnienia tradycyjnych metod wnioskowania statystycznego. Jednak dzięki procedurze *bootstrappingu* możliwym jest przewycięzenie tych mankamentów poprzez zapewnienie szacunkowych przedziałów ufności dla wskaźników efektywności uzyskanych za pomocą DEA. Jedną z takich metod (Simar i Wilson 1998, 2004) działa w następujący sposób. Oznaczmy wektor s wyników osiągniętych przez jednostkę j ($j=1, \dots, n$) jako y_j i wektor m nakładów używanych przez jednostkę j jako x_j . Etap 1 obejmuje oszacowanie wskaźników efektywności dla zbioru danych.

Wykorzystujemy metodę DEA przy danych nakładach i wynikach w celu oszacowania wskaźników efektywności oznaczonych jako: $\widehat{D}(x_j, y_j)$. Oszacowania te muszą znajdować się w obrębie jedynki, co zapewnia się przez obliczenie: $2 - \widehat{D}(x_j, y_j)$ dla każdego $\widehat{D}(x_j, y_j)$. $j=1, \dots, n$, dostarczając w sumie $2n$ obserwacji (n obserwacji dla $\widehat{D}(x_j, y_j)$ i n obserwacji dla $2 - \widehat{D}(x_j, y_j)$).

W drugim kroku wyznacza się wartości bootstrapowe, a krok ten złożony jest z wielu etapów. Po pierwsze, wyznaczone zostaje pasmo (ang. *bandwith*) h . Może ono zostać ustalone arbitralnie np. $h=0.05$ (z późniejszymi niewielkimi zmianami w celu oceny wpływu wielkości h na szacowany przedział ufności) albo może zostać wyznaczone według innych kryteriów (przykładowe są omówione w pracy Simara i Wilsona (1998)). Następnie losowana jest n -elementowa próbka niezależnych i identycznie rozłożonych obserwacji (zapisanych jako: $e_j, j=1, \dots, n$) z funkcji gęstości w formie rozkładu kernala. Po trzecie, losuje się n niezależnych wartości (oznaczonych jako $d_j, j=1, \dots, n$) ze zbioru $2n$ odzwierciedlających funkcje odległości. Na ich podstawie oblicza się średnią:

$$\bar{d} = \sum_{j=1}^n d_j / n \quad (1.42)$$

a potem oblicza się wartość:

$$d_j^* = \bar{d} + \sqrt{\left(1 + \frac{h^2}{s^2}\right)} (d_j + h e_j - \bar{d}) \quad (1.43)$$

gdzie s^2 jest wariancją próby dla \bar{d} . W końcu następuje obliczenie wartości bootstrapowej (D_j^*) jako:

$$D_j^* = \begin{cases} d_j^* & \text{jeżeli } d_j^* \leq 1 \\ 2 - d_j^* & \text{dla pozostałych} \end{cases} \quad (1.44)$$

W kroku 3 otrzymuje się bootstrapowe oszacowania wskaźnika efektywności. Posiadając próbę bootstrapową, określa się wektory nakładów i wyników (x_j^*, y_j^*) jako:

$$y_j^* = D_j^* y_j / \widehat{D}(x_j, y_j), \quad (1.45)$$

$$x_j^* = x_j, \quad (1.46)$$

Otrzymuje się daną liczbę np. B bootstrapowych wskaźników efektywności dla każdej jednostki j ($j=1, \dots, n$) poprzez zastosowanie metody DEA dla próby bootstrapowej B -razy. Te oszacowania bootstrapowe mogą być oznaczone dla jednostki k jako $\{\widehat{D}_b^*(x_k, y_k)\}_{b=1}^B$.

W czwartym kroku oblicza się przedziały ufności dla rzeczywistych wartości wskaźników efektywności. Przedział ufności $100(1-\alpha)\%$ dla wskaźnika efektywności jednostki k jest obliczany poprzez znalezienie wartości b_α, a_α jako:

$$\Pr(-b_\alpha \leq \widehat{D}(x_j, y_j) - D(x_j, y_j) \leq -a_\alpha) = 1 - \alpha. \quad (1.47)$$

Wartości b_α, a_α nie są znane, ale mogą zostać wyestymowane przy użyciu bootstrapowych szacowań: $\{\widehat{D}_b^*(x_k, y_k)\}_{h=1}^B$ poprzez uszeregowanie wartości $\widehat{D}_b^*(x_k, y_k) - \widehat{D}(x_j, y_j)$ od najmniejszej do największej i odcięcie $(100\alpha/2)\%$ obserwacji z każdego końca. Estymatory: $-b_\alpha$ i $-a_\alpha$ (oznaczone jako: $-\hat{b}_\alpha$ i: $-\hat{a}_\alpha$) są końcami pozostałego przedziału tak, że: $\hat{a}_\alpha \leq \hat{b}_\alpha$. Przybliżenie równania (1.47) wynosi:

$$\Pr(-\hat{b}_\alpha \leq \widehat{D}(x_j, y_j) - D(x_j, y_j) \leq -\hat{a}_\alpha) \approx 1 - \alpha, \quad (1.48)$$

A przedział ufności $100(1-\alpha)\%$ dla wskaźnika efektywności jednostki k jest znaleziony poprzez obliczenie:

$$[\widehat{D}(x_j, y_j) + \hat{a}_\alpha \widehat{D}(x_j, y_j) + \hat{b}_\alpha]. \quad (1.49)$$

Dodatkowe informacje dostarczone przez DEA. Warto na tym etapie rozważyć zalety i wady metody DEA w porównaniu do statystycznych metod parametrycznych w kontekście badań nad funkcjonowaniem instytucji edukacji. Powyżej zostało wykazane, że standardowa metoda DEA może dostarczać informacji na temat efektywności (technicznej i skali produkcji) dla wszystkich jednostek decyzyjnych w próbie. Ponadto kolejne procedury mogą dostarczyć informacji na temat efektywności alokacyjnej, a nawet efektywności wartości. Trzeba jednak zauważyć, że DEA oferuje nie tylko prostą miarę efektywności technicznej. Dzięki rozwiązaniu problemu DEA otrzymujemy bardzo dużo informacji odnośnie zarządzania jednostką, które mogą zostać wykorzystane do działań mających na celu zwiększenie efektywności jednostki. Będzie to rozwinięte w poniższej części rozdziału.

Po pierwsze, wyniki DEA mogą być wykorzystywane do identyfikacji dla każdej technicznie nieefektywnej jednostki w odniesieniu do jednostek efektywnych, które powinna naśladować. Są one określane jako jednostki referencyjne. Te ostatnie dla danej jednostki nieefektywnej nie muszą składać się ze wszystkich jednostek znajdujących się na granicy, a raczej obejmują producentów, którzy dominują w nieefektywnej jednostce decyzyjnej¹³. W prostym przykładzie zilustrowanym na Rysunku 1.8

¹³ Oznacza to, że liniowa kombinacja wektorów ich nakładów i wyników jest bardziej efektywna aniżeli w przypadku producentów nieefektywnych.

nieefektywna jednostka E mogłaby stać się efektywna, przesuając się do punktu E' znajdującego się na granicy produkcji, który jest kombinacją nakładów i wyników osiągniętych przez jednostki B i C . One właśnie są jednostkami referencyjnymi dla nieefektywnej jednostki E . Dla jednostki nieefektywnej wyznaczenie jednostek referencyjnych może być przydatne w podjęciu decyzji, które jednostki naśladować w celu zwiększenia efektywności.

Po drugie, granica efektywności może być wykorzystana do wyznaczenia celów do osiągnięcia przez jednostki nieefektywne (Färe i in. 1989; Pedraja-Chaparro i in. 1999). Punkty te mogą leżeć w dowolnym miejscu wzdłuż granicy efektywności. Przydatnym okazuje się często wyznaczenie celu jako wyników, które są konieczne, aby przybliżyć jednostkę do granicy efektywności technicznej, jeśli nakłady pozostają niezmienione (punkt E' na Rysunku 1.8 odzwierciedla poziom wyników, jakie nieefektywna jednostka E musi osiągnąć, aby stała się efektywna). Alternatywnie można nałożyć wymogi odnośnie poziomów nakładów bez zmian wyników. Możliwa jest dowolna kombinacja tych podejść. Tak ustalone cele są osiągalne i realistyczne w tym sensie, że odzwierciedlają one aktualnie najlepszą stosowaną praktykę. Celem jest wskazanie nieefektywnym jednostkom dokładnie tego, co mają zrobić, aby osiągnąć wyniki uzyskiwane przez najlepszych producentów. Nieefektywne jednostki mogą same na siebie nakładać wymogi lub mogą one zostać narzucone przez organizacje zewnętrzne.

Po trzecie, łatwość, z jaką DEA jest aplikowana do jednostek z wieloma nakładami i wynikami (Banker i in. 1984, 1989; Färe i in. 1989; McCarty i Yaisawarng 1993; Pedraja-Chaparro i in. 1999; Mante, 2001), jest poważnym atutem tej metody w kontekście badań instytucji edukacji, gdzie dochodzi do wytwarzania wielu wyników z wielu nakładów. Przykładowo uniwersytety prowadzą działalność dydaktyczną i badania naukowe. Wyniki ich działalności mogą być mierzone poprzez np. sukces absolwentów na rynku pracy (wynik działalności dydaktycznej), jak i przez książki, artykuły i prezentacje konferencyjne (wynik prowadzonych badań naukowych). Wyniki te są osiągane przy wykorzystaniu wielu nakładów np. studentów (dla których powinna być dostępna informacja o ich ilości, jak i jakości), pracowników i innych materiałów¹⁴.

Po czwarte, nieparametryczny charakter metody DEA pozwala każdej jednostce na wybór własnych wag dla nakładów i wyników, zamiast opierać się na ich relatywnym znaczeniu (Pedraja-Chaparro i in. 1999). Dla instytucji z sektora edukacji jest to duża wartość dodana, ponieważ dla tego sektora nie posiadamy najczęściej informacji o cenach, które mierzą względne znaczenie wyników i nakładów. Taka

¹⁴ Miary nakładów i wyników dla instytucji sektora edukacji będą analizowane bardziej szczegółowo w części 5.

elastyczność pozwala więc jednostce mieć różne cele w różnych warunkach (Pedraja-Chaparro i in. 1999).

Wreszcie, wszystkie warianty DEA mają wspólną cechę – efektywność jest mierzona w stosunku do jednostek wzorcowych o najlepszych praktykach. W modelach tych nie ma absolutnego, z góry narzuconego standardu efektywności. To podejście ma tę zaletę, że wiadomym jest, że osiągalna jest najlepsza praktyka stosowana przez jednostkę wzorcową. Nie można jednak zapominać o wadach metody DEA. Najprawdopodobniej największą z nich jest jej niestochastyczny charakter. Tak więc pominięte zmienne, błędy pomiaru nakładów i wyników oraz wszelkie inne zakłócenia statystyczne spowodują obciążenie wskaźników efektywności. W kontekście edukacji specyfikacja i pomiar zmiennych opisujących nakłady i wyniki działalności jednostek są szczególnie trudne (temat rozwinięty w części 5). Błędy w obliczonych wskaźników efektywności wynikające z pominięcia zmiennych, użycia próbki mniejszej niż populacja, związanych z obserwacjami odstającymi i błędów pomiaru są wysoce prawdopodobne, chociaż zakres problemów jest nieznan. Rozwój stochastycznej metody DEA (Lovell 1993) i wprowadzenie innych adaptacji DEA w celu zmniejszenia problemu błędów pomiarowych w zmiennych nakładów i wyników (Thanassoulis, 1999) są technikami niezmiernie ważnymi dla rozwoju metody DEA. Pozostałe problemy związane z funkcjonowaniem tej metody wynikają z wcześniej opisanych zalet. Chociaż DEA opiera się na wielu nakładach i wynikach, to ich specyfikacja, w szczególności ich liczba, może wpływać na wartość wskaźników efektywności (Pedraja-Chaparro i in. 1999). Ponadto ponieważ DEA ma charakter niestatystyczny i nieparametryczny, nie ma znanych testów parametrycznych na podstawie których można sprawdzić poprawność modelu. Dlatego zaleca się rozważenie wyników uzyskanych za pomocą różnych specyfikacji w celu oceny wrażliwości wyników na ich zmiany. Po drugie, duża liczba wyników i/lub nakładów może prowadzić do współliniowości, której wpływ na wyniki DEA jest stosunkowo niedostateczny rozpoznany. Będzie to tematem części 5. Po trzecie, założenie VRS, w którym faktycznie nie ma efektów skali, może wpływać na wskaźniki efektywności. Jak zostało pokazane na Rysunku 1.9, granica VRS obejmuje dane bardziej przyległe aniżeli granica CRS. W konsekwencji jeżeli zostanie zastosowany model VRS, a prawdziwym modelem jest CRS, to wskaźniki efektywności dla jednostek ekstremalnych (najmniejszych i największych jednostek) będą przeszacowane (Dyson i in. 2001).

4. Techniki pomiaru efektywności: analiza na poziomie indywidualnym

Istnieją różne argumenty krytyczne w stosunku do metod, w których pomiar jednostek produkcyjnych jest dokonywany na danych zagregowanych. Woodhouse i Goldstein (1988) podają argumenty przeciw stosowaniu reszt z modeli regresji

do użycia dla danych zagregowanych jako pomiaru efektywności jednostek takich jak szkoły, lokalne jednostki nadzoru edukacji (ang. LEA) czy uniwersytety. Argumenty te mogą zostać podniesione w stosunku do wszystkich statystycznych metod pomiaru efektywności.

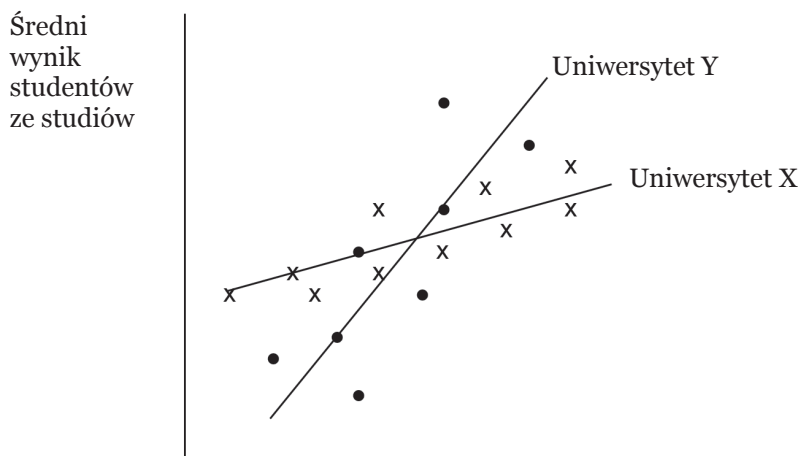
Po pierwsze, wykorzystanie danych mierzonych na poziomie organizacji nie pozwala na wychwycenie zmienności relacji wewnątrz jednostki. Na przykład w analizie uniwersytetów wyniki działalności mogą zostać ukazane jako średni wynik osiągnięty przez wszystkich studentów, a nakłady jako średnia ilość punktów uzyskanych w czasie rekrutacji. Analiza na poziomie uniwersytetu nakłada taką samą zależność pomiędzy uzyskiwanymi średnimi wynikami a ponoszonymi średnimi nakładami dla wszystkich uniwersytetów. Uniwersytet X może uzyskiwać wyższe wyniki niż uniwersytet Y dla studentów o niskich punktach na wejściu, ale sytuacja może wyglądać odwrotnie dla studentów o wysokich punktach na wejściu (zob. Rysunek 1.11). Idealnie byłoby, gdyby zależności pomiędzy wynikami a nakładami, jakie istnieją wewnątrz uniwersytetu (jeżeli są znane), były wzięte pod uwagę podczas mierzenia efektywności. Dotychczasowe badania oparte o dane zagregowane nie miały takiej możliwości¹⁵.

Po drugie, reszty z modeli regresji mogą się znacznie różnić w zależności od zmiennych uwzględnionych w modelu. Zmiana reszty może powodować, że rankingi jednostek będą się różnić w zależności od różnych modeli statystycznych. Ponadto wybór konkretnego modelu może być trudny z uwagi na to, że to, co bierzemy za dobre odzwierciedlanie danych na jego gruncie, jest najczęściej ograniczane przez małą liczebność próby.

Pomiar efektywności jednostek za pomocą DEA na poziomie zagregowanym budzi poważne zastrzeżenia (Woodhouse i Goldstein 1988). Wynik wskaźnika efektywności dla jednostki jest obliczany jako stosunek ważonych wyników do ważonych nakładów. W prostym przykładzie, gdzie mamy jeden wynik (y) i jeden nakład (x) (y – średni z uzyskanych rezultatów, a x – średnia z uzyskanej ilości punktów na wejściu np. średnia ilość punktów z rekrutacji) efektywność jednostki k jest mierzona jako iloraz: $\theta_k = y_k / x_k$. Jakkolwiek istnieje związek pomiędzy wynikami uzyskanymi na koniec studiów a nakładami na wejściu np. punktami z rekrutacji, zakładając, że zależność ta ma formę $y_k = a + bx_k$, to efektywność jednostki k jest wyrażona jako: $\theta_k = a/x_k + b$ i jest odwrotnie proporcjonalna do miar nakładów. Przykład ten może zostać rozszerzony na sytuacje z wieloma nakładami i wieloma wynikami. Dlatego DEA przy użyciu danych zagregowanych może mierzyć efektywność jednostki w sposób niedokładny.

¹⁵ Thanassoulis (1996) rozwinął metodę opartą o DEA w celu wyznaczenia odpowiednich celów dla szkół o różnych efektywnościach tak, aby zidentyfikować jednostki wzorcowe do naśladowania przez szkoły nieefektywne.

Rysunek 1.11 Efektywność dwóch uniwersytetów



Średnia punktacja studentów uzyskana przy przyjęciu na studia

Dlatego w tej sekcji omówiono niektóre metody mierzenia efektywności organizacji wykorzystujących dane dla indywidualnych jednostek wchodzących w jej skład.

Statystyczne metody parametryczne. Istnieje wiele metod statystycznych, które mogą zostać zastosowane w stosunku do indywidualnych uczniów i studentów w celu uzyskania oceny efektywności instytucji, do której uczęszczają. To jaka metoda będzie zastosowana, zależy od natury zmiennej zależnej. Tak więc regresja MNK jest odpowiednia, jeśli zmienna zależna jest zmienną ciągłą (na przykład zarobki absolwentów), modele logitowe lub probitowe są właściwe dla zmiennych binarnych (np. czy student ukończył studia), a uporządkowane modele logitowe lub probitowe mogą być stosowane, gdy zmienna zależna jest zmienną przyjmującą jedną ze skończonych uporządkowanych liczby możliwości (np. wynik uzyskany przez absolwenta na dyplomie albo wynik egzaminu ucznia). Modele statystyczne, które wykorzystują dane na poziomie indywidualnych jednostek w celu oceny związku między wynikami procesów kształcenia a nakładem, ale które nie uwzględniają efektów instytucjonalnych, mogą być uznane za niezadowolające z dwóch powodów. Po pierwsze, statystyczne testy istotności mogą być obciążone i przeszacowane. Po drugie, jeśli efekty instytucjonalne nie są wzięte pod uwagę, model nie mówi nic na temat wpływu instytucji na wyniki uzyskiwane przez uczestników procesu edukacji (Goldstein 1997); oznacza to, że taki model nie mówi nic na temat efektywności samej instytucji.

Jednak modele wielopoziomowe (ang. *multi level*) uwzględniają efekty instytucjonalne w ocenie związku pomiędzy indywidualnymi nakładami i wynikami.

W modelach tych zakłada się, że dane do analizy są hierarchiczne; na przykład uczniowie są częścią szkoły, do której chodzą, a studenci uniwersytetów, do których uczęszczają. Są to przykłady danych pochodzących z dwóch poziomów: studenci (uczniowie) to poziom 1, a uniwersytety (szkoły) to poziom 2. Można dodać dodatkowe poziomy do analizy; na przykład uczniowie są częścią szkoły, a ta podlega lokalnym władzom oświaty. Poniżej przedstawiony zostanie krótki przegląd metod wielopoziomowych ze szczególnym uwzględnieniem jego zastosowania w pomiarze efektywności. Zainteresowanych czytelników kierujemy do prac: Aitkina i Longforda (1986) i Goldsteina (1995), gdzie zamieszczone są bardziej szczegółowe informacje oraz dodatkowe rozszerzenia. Na początku rozważmy prosty model w kontekście szkolnictwa wyższego. Niech y_{ij} oznacza wynik osiągnięty przez i -tego studenta uczęszczającego na j -ty uniwersytet, model zapisać można jako:

$$y_{ij} = \beta_j + e_{ij} = \beta_0 + u_j + e_{ij} \quad (1.50)$$

Wynik uzyskiwany przez indywidualnego studenta może zostać rozbity na związany z jednostką do której uczęszcza (β_j) oraz odchyleniem od wpływu instytucji (e_{ij}). Wkład uniwersytetu (β_j) może być dalej podzielony na wartość średnią dla wszystkich uniwersytetów (β_0) oraz na odchylenie od średniej (u_j). Dlatego u_j jest często określane jako efekt (efekt uniwersytetu) i zostanie omówione poniżej.

Należy przyjąć kilka założeń. Najpierw zakłada się, że uniwersytety są losową próbą pochodzącą z całej populacji uniwersytetów. Składnik u_j wśród uniwersytetów ma rozkład normalny o średniej zero i wariancji σ_u^2 . Reszty dla indywidualnych studentów (e_{ij}) też mają rozkład normalny o średniej zero i wariancji σ_e^2 . Odpowiednie programy komputerowe są w stanie oszacować nieznanne parametry β_0 , σ_u^2 , σ_e^2 na podstawie, których można obliczyć u_j (zapisany jako \hat{u}_j) przy użyciu poniższego równania:

$$\hat{u}_j = \frac{n_j \sigma_u^2}{n_j \sigma_u^2 + \sigma_e^2} \cdot \frac{\sum_i (y_{ij} - \hat{y}_{ij})}{n_j}, \quad (1.51)$$

gdzie n_j – liczba studentów na uniwersytecie j . Efekt dla każdego uniwersytetu \hat{u}_j ma obliczony błąd próbkowy na podstawie, którego można obliczyć przedziały ufności. Na podstawie oszacowań i przedziałów ufności można ocenić działalność uniwersytetów. Instytucje szereguje się najczęściej według wartości \hat{u}_j od najmniejszej do największej z zaznaczeniem dolnej i górnej granicy przedziału ufności. Przyjmuje się, że jeżeli przedziały się nakładają, to brak jest zróżnicowania w działalności instytucji.

Prosty model może zostać rozszerzony poprzez włączenie kolejnych determinant zmiennej zależnej. Na przykład najlepszym wyznacznikiem osiągnięć studentów uzyskanych na dyplomie jest wynik uzyskany w momencie rekrutacji (x_{ij}). Model można teraz zapisać jako:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij} \quad . \quad (1.52)$$

W modelu tym współczynnik dla zależności pomiędzy y_{ij} i x_{ij} jest stały, podczas gdy punkt przecięcia jest zmienny dla uniwersytetów. Dlatego parametry β_0 i β_1 są stałe, a u_j i e_{ij} są losowymi składowymi modelu. Dodanie zmiennej objaśniającej prowadzi do interpretacji reszty \hat{u}_j jako wskaźnika efektywności uniwersytetu w kontekście wartości przez niego dodanej. Odpowiednio można dodać kolejne zmienne objaśniające do modelu. Na końcu przedstawiamy model, w którym zarówno stała, jak i nachylenie pomiędzy y_{ij} i x_{ij} mogą być zmienne dla uniwersytetów, co można zapisać jako:

$$y_{ij} = \beta_0 + \beta_{1j} x_{ij} + u_j + e_{ij}, \quad (1.53a)$$

$$\beta_{1j} = \beta_1 + v_j \quad . \quad (1.53b)$$

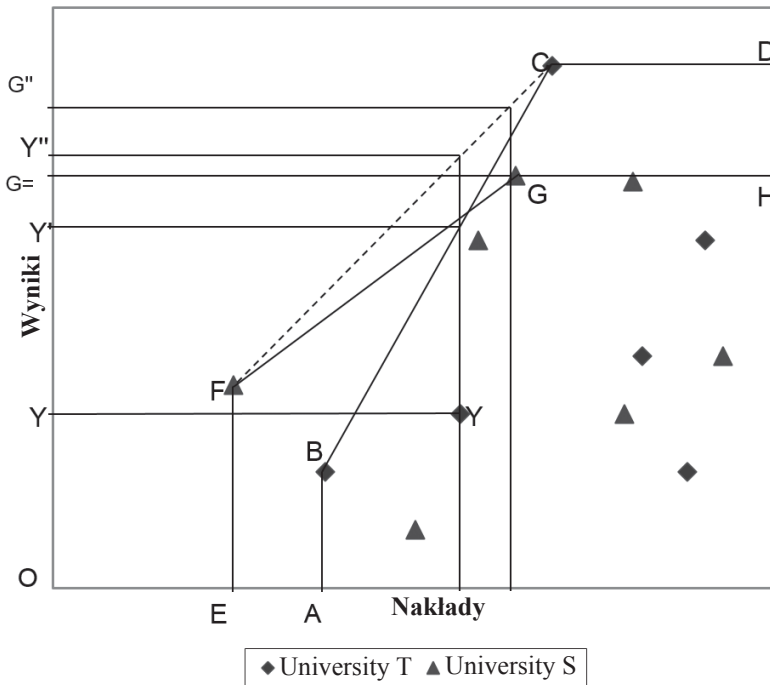
Wynika z tego, że średnie nachylenie dla populacji uniwersytetów jest równe parametrowi β_1 , a odchylenie dla poszczególnych uniwersytetów wynosi v_j . Składniki u_j i v_j posiadają wielowymiarowy rozkład normalny (tutaj jest to dwuwymiarowy rozkład normalny, gdyż są tylko dwie zmienne losowe na poziomie drugim) o wariancji oczekiwanej równej zero. Wariancja u_j mierzy odchylenie dla stałych, punktów przecięcia $\text{var}(u_j) = \sigma_u^2$; a wariancja v_j zmienność nachylenia i wynosi $\text{var}(v_j) = \sigma_v^2$. Kowariancja pomiędzy u_j i v_j odpowiada kowariancji pomiędzy stałą a nachyleniem i wynosi $\text{cov}(u_j, v_j) = \sigma_{u,v}$. Wyniki dla poszczególnych studentów odchylają się od linii o e_{ij} . W modelu tym parametry β_0 i β_1 są stałe, a v_j , u_j i e_{ij} są współczynnikami losowymi.

Problemem modeli wielopoziomowych jest to, że w celu modelowania złożonych procesów edukacji same muszą być bardzo skomplikowane, przez co ich interpretacja staje się utrudniona. Ponadto bardziej złożone modele (na przykład te z udziałem nieciągłych zmiennych zależnych) mogą być trudne do wyliczenia.

Metody niestatystyczne i nieparametryczne. Alternatywą dla metod statystycznych jest DEA, gdzie jednostka decyzyjna, zamiast być szkołą, uniwersytetem, wydziałem lub regionem, jest w rzeczywistości indywidualnym uczniem lub studentem

(Thanassoulis 1999; Portela i Thanassoulis 2001; Thanassoulis i Portela 2002). Rozważyć można, dla przykładu, sytuację, w której rezultatami są wyniki uzyskiwane przez poszczególnych studentów, a nakładami są ich kwalifikacje wyrażone jako wynik punktów uzyskanych w momencie przyjęcia na studia. Dla każdego studenta można obliczyć jego efektywność poprzez zastosowanie metody DEA, ale wynik ten będzie miał w sobie dwa komponenty: jeden związany z wysiłkami studentów, a drugi odpowiadający efektywności nauczycieli, którzy go uczyli. Aby ocenić efektywność instytucji szkolnictwa wyższego, trzeba na początku rozbić wskaźniki efektywności studentów na te dwie składowe – metoda taka została zaproponowana przez Portelę i Thanassoulisa (2001, 2002). Rozważmy hipotetyczny zestaw danych dla studentów z dwóch uniwersytetów. Każdy z nich produkuje absolwentów, których jakość jest mierzona za pomocą wyników uzyskanych na dyplomie przy użyciu początkowych umiejętności studenta mierzonych punktami z rekrutacji. Dane na temat nakładów i wyników dla wszystkich studentów zostały zobrazowane na Rysunku 1.12.

Rysunek 1.12 Dekompozycja efektywności za pomocą DEA



Granica *EFCD* obejmuje wszystkich studentów i można ją nazwać granicą efektywności studenta w stosunku do wszystkich instytucji. Studenci leżący w obszarze *EF* i *CD* znajdują się na granicy, ale nie są efektywni (z powodu luzów). Stosując tradycyjną definicję efektywności DEA (zorientowaną na wyniki), student *F*, który leży na granicy ma wskaźnik efektywności równy 1, natomiast student *Y*, który znajduje się wewnątrz granicy, ma ogólny poziom efektywności równy OY/OY'' , który jest mniejszy niż 1. Innymi słowy, OY/OY'' oznacza stosunek wyników uzyskanych przez studenta *Y* do najlepszych osiągnięć uzyskanych w grupie studentów ze wszystkich uniwersytetów przy wzięciu pod uwagę kwalifikacji na wejściu studenta *Y*.

Jednakże wskaźnik efektywności wyznaczony przez porównanie osiągnięć studentów ze wszystkich uniwersytetów nie bierze pod uwagę wpływu danej instytucji na wyniki osiągane przez studentów. Przykładowo studenci z uniwersytetu *T* mają swoją własną granicę (nazywaną granicą efektywności studenta w ramach własnej instytucji) wyznaczoną przez *ABCD*. Podobnie granica dla uniwersytetu *S* wynosi *EFGH*. Dlatego student *Y* (z uniwersytetu *T*) osiąga efektywność w stosunku do własnej instytucji w wysokości OY/OY' , która oznacza stopień wyników uzyskanych przez studenta w odniesieniu do najlepszego wyniku uzyskanego przez studentów z jego uczelni, biorąc pod uwagę jego wejściowe kwalifikacje. Odległość $Y'Y''$ mierzy wpływ uniwersytetu na wyniki osiągane przez studenta *Y*. Ocena efektywności pomiędzy instytucjami dla studenta *Y* wynosi OY'/OY'' i jest zmienna w zależności od nakładów.

Podsumowując, student *Y* charakteryzuje się wskaźnikiem efektywności OY/OY'' , z czego OY/OY' jest wynikiem jego własnych wysiłków, a OY'/OY'' jest konsekwencją efektywności uniwersytetu, na który uczęszcza. Można go skonstrastować ze studentem *G* z uniwersytetu *S*, który charakteryzuje się wskaźnikiem efektywności w ramach własnej instytucji równym 1 (student osiąga najlepsze wyniki w porównaniu do studentów z tej samej uczelni, biorąc pod uwagę jego kwalifikacje na wejściu), ale jego wskaźnik efektywności w ramach wszystkich instytucji w wysokości $OG'/OG'' = OG/OG''$ jest mniejszy niż 1, ponieważ jego własny uniwersytet nie jest efektywny.

Efektywność każdej z instytucji może zostać zbadana poprzez porównanie wskaźników efektywności uczelni ze wskaźnikami efektywności własnych studentów. Komponent ten jest miarą efektywności samej instytucji i nie zawiera w sobie efektów związanych z różnymi wysiłkami studentów. Porównanie wszystkich trzech składników może być przydatne, w szczególności dla zarządzających organizacjami w celu określenia o ile można zwiększyć efektywność danej instytucji (Portela i Thanassoulis 2001; Thanassoulis i Portela 2002).

5. Badania empiryczne

Dokładne omówienie wszystkich empirycznych zastosowań metod opisanych w poprzednich częściach w kontekście edukacji i szkolnictwa wyższego wykracza poza zakres niniejszego opracowania. Dlatego też odnosimy się do ograniczonej liczby badań w celu uwypuklenia niektórych problemów i zastosowań technik pomiaru efektywności w kontekście edukacji. Bardziej kompletne podsumowanie wyników badań empirycznych można znaleźć w aneksie do niniejszego tekstu¹⁶.

Jednostka analizy

W studiach nad efektywnością wykorzystuje się rozmaite jednostki analizy. W ramach sektora szkolnego taką jednostką może być uczeń, szkoła, region i lokalne władze odpowiedzialne za edukację (Goldstein i Thomas 1996; Bradley i in. 2001; Engert 1996; Mayston i Jesson 1988). W przypadku szkolnictwa wyższego dotychczasowe badania starały się ocenić jego działanie, posługując się jako jednostką analizy pojedynczymi studentami, wydziałami tego samego typu w ramach różnych uniwersytetów, centrami, w których generowane są koszty w ramach różnych uniwersytetów, różnymi wydziałami w ramach jednego uniwersytetu, jak również samymi uniwersytetami (Johnes 1996, 2003; Beasley 1995; Friedman i Sinuany-Stern 1997). Badania prowadzone na poziomie indywidualnym dostarczają znaczną ilość informacji, ale ich rezultaty mogą być skomplikowane i trudne w interpretacji. Z drugiej strony badania prowadzone na poziomie danych zagregowanych są na ogół proste w interpretacji, co osiąga się jednak kosztem utraty informacji w procesie agregacji. Być może należy postawić pytanie: czy analizy prowadzone na różnych poziomach danych dostarczają takich samych, czy różnych miar efektywności?

Zwolennicy i osoby stosujące modelowanie wielopoziomowe twierdzą, że wyłącznie zastosowanie tego podejścia do danych z indywidualnego poziomu jest jedyną metodą, która jest zdolna właściwie uchwycić złożone efekty poszczególnych instytucji wpływających na wyniki studentów (Aitkin i Longford 1986). Pokazano, że wnioski na temat efektów działalności szkół uzyskane za pomocą modelowania wielopoziomowego prowadzą do innych rankingów szkół niż te, które są uzyskane za pomocą wykorzystania danych zagregowanych (Sammons i in. 1993). Podobne wyniki osiąga się przy zastosowaniu modeli regresji probitowej zastosowanych do danych na temat studentów i absolwentów (Baratti 2002; Naylor i in. 2000; Smith i Naylor 2001a; Smith i in. 2000). Dotychczas jednak żadne z badań nie porównywało ze sobą rankingów instytucji opartych na efektach modelowania

¹⁶ Z powodów ograniczonej liczby stron czasopisma odsyłamy zainteresowanych czytelników do oryginalnego tekstu (przyp. tłum.).

wielopoziomowego, efektach marginalnych z modelu probitowego czy rankingu opartego na danych zagregowanych.

Posługiwanie się metodą DEA w przypadku indywidualnych obserwacji, a nie całych instytucji, jest stosunkowo niedawno wprowadzoną innowacją. Pomimo sugestii, że zastosowanie DEA do poszczególnych uczniów może być pożyteczne pod kątem dostarczania nauczycielom informacji, którzy uczniowie nieefektywnie wykorzystują zasoby (Bessent i in. 1982), większość badań prowadzonych na polu edukacji wykorzystuje dane na temat wydziałów, instytucji czy poziomu administracyjnego. Bez wątplenia jest to związane z ogromnymi wymogami obliczeniowymi, które są potrzebne przy zastosowaniu metody DEA na poziomie indywidualnym: liczba obserwacji warunkuje zarówno liczbę i złożoność programowania liniowego, które muszą zostać rozwiązane. Dotychczas metoda DEA była wykorzystana na poziomie indywidualnym do bardzo różnej liczby obserwacji i obejmowała od ponad 2500 jednostek (Johnes 2003) do ponad 6500 jednostek (Portela i Thanassoulis 2001). Wstępne ustalenia sugerują, że rankingi uniwersytetów opartych na metodzie DEA zastosowanej do poziomu indywidualnego różnią się od tych, które powstały w oparciu o metodę DEA zastosowaną do tych samych danych, ale zagregowanych do poziomu uniwersytetu (Johnes 2003). Z pewnością potrzeba więcej badań w tym zakresie, dzięki którym będzie można wyznaczyć kierunki badań nad efektywnością jednostek sektora edukacyjnego.

Pytanie, które jest być może istotniejsze w kontekście szkolnictwa wyższego niż szkół, dotyczy tego, czy dane powinny być agregowane na poziomie różnych dyscyplin. Istnieją dowody na to, że czynniki determinujące wyniki prowadzonej działalności badawczej przez szkoły wyższe, oceny uzyskiwane przez studentów oraz procent osób, które nie kończą studiów, różnią się w zależności od wydziału i reprezentowanej dyscypliny. Świadczy to o tym, że badania efektywności w kontekście uniwersytetów powinny być przynajmniej zdezagregowane na poziomie dyscyplin i podaje się w wątpliwość sensowność porównań różnych wydziałów w ramach jednego uniwersytetu.

Na poziomie szkół istnieją pewne dowody na to, że rola czynników związanych z pochodzeniem uczniów czy ich początkowymi umiejętnościami odgrywa większą rolę w wyjaśnianiu różnic w umiejętności czytania niż w zakresie zdolności matematycznych i że wpływ działalności szkoły jest większy na osiągnięcia z matematyki niż z czytania (Sammons i in. 1993). Co więcej, korelacja rankingów szkół opartych na dwóch modelach efektów wytwarzanych przez szkołę – jeden wyprowadzony z wyników w umiejętności czytania, drugi z wyników w matematyce – nie jest szczególnie wysoka (Sammons i in. 1993). Może to zatem świadczyć o tym, że nawet na poziomie szkoły ewaluując funkcjonowanie instytucji, powinno przykładać się wagę do badania różnic pomiędzy wykładanymi przedmiotami.

Wyniki

Podobnie jak w przypadku poziomym, na którym prowadzona jest analiza efektywności, na specyfikę wyników również oddziałują względy dotyczące cech danych. W sytuacji idealnej powinno się uwzględnić miary wszystkich wyników działalności instytucji. Problem (poza ograniczeniami samych danych) polega na tym, że nie jest do końca jasne, czym są wyniki procesu edukacyjnego. Wynik prowadzonej działalności dydaktycznej może zostać podzielony na szereg składników, z których wszystkie można zaklasyfikować jako rezultaty procesu edukacyjnego. Wśród nich wymienić można: osiągnięcia studentów, wzrost produktywności siły roboczej, obecne i przyszłe dodatkowe korzyści konsumpcyjne.

Osiągnięcia studentów często udaje się przybliżyć, gdyż ocenianie odbywa się na wielu szczeblach edukacji, a tym samym często dostępne są dane na temat wyników różnych testów i egzaminów. Dane na temat liczby uczniów/studentów kończących naukę albo uzyskujących daną ocenę są często używane jako miary kształcenia w zagregowanych analizach szkół (Bessent i Bessent 1980; Kirjavainen i Loikkanen 1998) i zagregowanych badaniach uniwersytetów (Athanassopoulos i Shale 1997; El-Mahgary i Lahdelma 1995; Engert 1996; Madden i in. 1997). Miary te nie biorą jednak w wystarczający sposób pod uwagę jakości osiągnięć, tym samym średnie wyniki z egzaminów uzyskane w danej instytucji i procentowy wskaźnik sukcesu są często preferowane w analizach na poziomie zagregowanym, a zdobyte punkty na egzaminie są oczywistą miarą wyników w analizach prowadzonych na poziomie indywidualnym (zob. np. Chalos i Cherian 1995; Bradley i in. 2001; Johnes i Taylor 1990 w kwestii badania przy wykorzystaniu danych na poziomie zagregowanym dla szkół podstawowych, gimnazjów i uniwersytetów; zob. Goldstein i Spiegelhalter 1996; Smith i Naylor 2001b na temat badań dla szkół gimnazjalnych i szkół wyższych prowadzonych przy użyciu danych na poziomie indywidualnym).

Pojawiają się wątpliwości, co do używania danych na temat wyników z egzaminów jako rezultatów działalności dydaktycznej, ponieważ egzaminy nie weryfikują wszystkich umiejętności, które mogą być rozpatrywane jako pożądane wyniki edukacji (Bifulco i Bretschneider 2001), a osiągnięcia akademickie są tylko jednym wynikiem procesu nauczania (Bessent i in. 1982). Miary efektywności odzwierciedlające wpływ instytucji na zwiększenie produktywności absolwentów na rynku pracy zawierałyby ten dodatkowy rezultat działalności dydaktycznej, jednak dane dotyczące sukcesu absolwentów na rynku pracy nie zawsze są dostępne. Pierwsza ankieta dotycząca ekonomicznych losów absolwentów w Wielkiej Brytanii (ang. *The First Destinations Survey*), która dostarczała informacji na temat losów absolwentów po sześciu miesiącach od zakończenia studiów, pozwoliła na stworzenie miar opartych na wynikach, które odzwierciedlają sukces na rynku pracy w badaniach efektywności brytyjskich uniwersytetów (Johnes i Taylor 1989b, 1990; Johnes 1996; Bratti i in. 2004; Smith

i in. 2000). Kolejne bazy danych pozwoliły również na stworzenie miar odnoszących się do zarobków uzyskiwanych przez absolwentów (Naylor i in. 2000, 2002; Belfield i Fielding 2001).

W procesie nauczania instytucje edukacyjne produkują również jednostki, które ponoszą porażkę w procesie nabywania kwalifikacji. Marnotrawstwo spowodowane porażką w kontekście egzaminów, absencji (na poziomie edukacji szkolnej) czy niekończenia edukacji (w przypadku szkół średnich czy szkolnictwa wyższego) są efektami ubocznymi procesu nauczania, a ich zakres jest często niewidoczny, np. jeśli używa się liczby absolwentów, liczby przyjętych studentów lub miar sukcesu na rynku pracy absolwentów jako sposobu odzwierciedlenia wyników nauczania. Dlatego w badaniach szkół i szkolnictwa wyższego do analiz włącza się też procent osób, które nie dokończyły edukacji lub procent obecności, by oddać ten aspekt związany z działalnością dydaktyczną (Bradley i in. 2001; Harrison i Rouse 2002; Ruggiero 1996; Johnes i Taylor 1989a; Johnes 1996; Mante 2001; Ramanathan 2001).

Ograniczenia związane z dostępnością danych na temat miar osiągnięć, sukcesu na rynku pracy oraz procent osób niekończących edukację sprawiły, że alternatywnie używa się danych dotyczących nakładów w celu odzwierciedlenia poziomu wyników. Przyjmując, że celem badań nad efektywnością jest ustalenie relacji pomiędzy nakładami a wynikami w celu oceny funkcjonowania instytucji, istnieje duża liczba analiz, w których wyniki działalności dydaktycznej są przybliżane przez miary nakładów takie jak całkowita liczba studentów czy liczba studentów nowoprzyjętych (Abbott i Doucouliagos 2003; Ahn i Seiford 1993; Avkiran 2001; Beasley 1990, 1995; Tomkins i Green 1988) czy liczba godzin prowadzonych zajęć dydaktycznych (Ahn i in. 1989; Friedman i Sinuany-Stern 1997).

Miary dodatkowych korzyści konsumpcyjnych (obecnych lub przyszłych) wynikające z edukacji czy innych efektów zewnętrznych procesu nauczania (takich jak rozwój pożądaných nawyków pracowniczych, korzyści dla firm w postaci ułatwienia procesu rekrutacyjnego i doradztwa zawodowego dla studentów) są trudne do skonstruowania. W żadnych z dotychczasowych badań miary tego typu nie zostały w sposób satysfakcjonujący włączone do analizy jako odzwierciedlenie wyników działalności instytucji.

Wyniki działalności instytucji szkolnictwa wyższego są zarówno związane z procesem kształcenia, jak i prowadzenia badań naukowych, tym samym analiza efektywności uniwersytetów nie byłaby kompletna bez wzięcia pod uwagę wydajności w produkowaniu wyników w kontekście działalności badawczej. Podobnie jak z wynikami kształcenia miary oparte na nakładach, takie jak przychód z badań czy wydatki na badania, są popularnymi miarami, które zastępują wyniki prowadzonej działalności badawczej (Beasley 1990, 1995; Friedman i Sinuany-Stern 1997; Johnes i Johnes 1993; Tomkins i Green 1988; Hashimoto i Cohn 1997). Bardziej satysfakcjonujące

miary wyników związane z prowadzoną działalnością badawczą, uwzględniające zarówno ilość, jak i jakość, oparte są o wskaźniki, do których informacje dostarczane są w procesie oceny eksperckiej. Dane te są dostępne obecnie dla uniwersytetów z Australii i Wielkiej Brytanii (Abbott i Doucouliagos 2003; Avkiran 2001; Sarrico i Dyson 2000; Tomkins i Green 1988; Johnes i Taylor 1990, 1992; Johnes 1996; Johnes i in. 1993). W obliczu braku danych z ocen eksperckich alternatywną miarę produktywności naukowej stanowi liczba publikacji (Johnes i Johnes 1992, 1993; Madden i in. 1997; Tomkins i Green 1988) lub np. liczba publikacji w czołowych czasopismach naukowych, która ma uwzględniać zarówno ilości, jak i jakości (Johnes i Johnes 1992, 1993; Madden i in. 1997).

Problemy w badaniach efektywności związane są często z opomiarowaniem rezultatów prowadzonej działalności. Generalnie zakłada się, że w ramach procesu edukacyjnego z wielu nakładów produkuje się wiele wyników. Jednak brakuje jednej spójnej metody mierzenia efektywności przy braniu pod uwagę zarówno wielkości nakładów, jak i wielkości wyników. Atrakcyjnym rozwiązaniem tego problemu jest metoda DEA, gdyż łatwo radzi sobie z obiema tymi kwestiami. Niewiele jednak wiadomo o tym, jak na wskaźniki efektywności uzyskiwane metodą DEA wpływa (jeśli w ogóle) współliniowość pomiędzy wynikami i sama natura produkcji (np. rezultaty produkowane mogą być łącznie, symultanicznie lub osobno).

Z drugiej strony metody statystyczne są zwykle stosowane, gdy mamy jedną zmienną zależną i wiele zmiennych objaśniających. Dlatego też metody te mogą z łatwością poradzić sobie z sytuacjami, w których mamy do czynienia z jednym nakładem i wieloma wynikami lub wieloma nakładami i jednym wynikiem. W stosunku do sektora edukacji z typowym dla niego procesem, w którym z wielu nakładów uzyskuje się wiele wyników, metody statystyczne nie są satysfakcjonujące¹⁷. Najczęściej spotykanym podejściem używanym w obliczaniach efektywności, gdy mamy do czynienia z wieloma nakładami i wieloma wynikami przy użyciu metod statystycznych, jest obliczanie osobnego równania dla każdego z wyników (Gray i in. 1984; Johnes i Taylor 1990; Yang i in. 1999; Johnes 1996; Sammons i in. 1993; Goldstein 1997; Goldstein i in. 1993). Alternatywnie oblicza się złożone miary wyników i używa się ich jako zmiennej zależnej (Sengupta i Sfeir 1986; Bates 1997). Podejście to może jednak powodować liczne problemy.

Po pierwsze, jeśli wyniki są produktem łącznym, efektywność może nie być poprawnie oszacowana za pomocą oddzielnych równań dla każdego z wyników, co powoduje problem polegający na niewzięciu pod uwagę możliwych zmian pomiędzy uzyskiwanymi wynikami działalności edukacyjnej pod wpływem różnej relokacji nakładów. Tym samym instytucja może zostać uznana za nieefektywną na podstawie

¹⁷ Stochastyczna metoda graniczna może zostać jednak zaadaptowana do oceny procesów z użyciem wielu nakładów i wielu wyników. Zob. Sickles i in. (2002), gdzie zastosowano ją w sektorze lotniczym.

oceny tylko jednego z uzyskiwanych wyników z sugestią zmiany nakładów w celu wzrostu wydajności. Z kolei w przypadku oddzielnych obliczeń brak jest informacji na temat tego, jak zmiana w nakładach przeznaczonych na poprawienie jednego z wyników oddziałuje (pozytywnie lub negatywnie) na inne wyniki procesu produkcyjnego (Maystone i Jesson 1998). Chizmar i Zak (1983) stwierdzają, że modelowanie wielu wyników w formie produktu łącznego jest lepszym rozwiązaniem pod względem empirycznym i teoretycznym od modelowania ich przy użyciu oddzielnych równań. Po drugie, jeśli otrzyma się wiele oddzielnych miar efektywności, to powstają problemy w ich interpretacji. Powyższe problemy mogą zostać rozwiązane np. za pomocą analizy głównych składowych (Johnes 1996).

Nakłady

Nakład pracy i kapitału ludzkiego wykorzystywanego w procesie edukacji często mierzy się przez liczbę studentów i kadry. Dane tego typu powinny być bez problemu dostępne na poziomie szkoły, obszaru administracyjnego czy szkolnictwa wyższego (zob. np. Ahn i Seiford 1993; Athanassopoulos i Shale 1997; Chalos 1997; Barrow 1991; Avkiran 2001; Grosskopf i Moutray 2001). Używano także bardziej wyrafinowanych miar nakładów pracy takich jak czas poświęcany przez uczniów na odrobienie pracy domowej czy liczba odbytych zajęć (Lovell i in. 1994), takie dane są jednak rzadko dostępne. Alternatywą dla uwzględniania danych zarówno o liczbie studentów, jak i kadry jest łączenie ich w jedną miarę złożoną np. stosunek studentów do kadry (Bradley i in. 2001; Belfield i Fielding 2001; Johnes i Taylor 1990; Mancebon i Mar Molinero 2000; Mante i O'Brien 2002).

Próby zmierzenia jakości nakładów osobowych polegają zazwyczaj na włączaniu do analizy informacji na temat zarobków pracowników lub zmiennych obrazujących kwalifikacje akademickie, wiek i doświadczenie pracowników (zob. np. Sengupta i Sfeir 1986; Ahn i Seiford 1993; Chalos 1997; Friedman i Sinuany-Stern 1997; Ruggiero 1996, 2000; Bradley i in. 2001; Soteriou i in. 1998). Jakość studentów ukazywana jest w badaniach efektywności prowadzonych na poziomie zagregowanym za pomocą średniej z wyników osiągnięć w populacji studentów zaczynających naukę lub proporcji osób z daną kwalifikacją w populacji studentów zaczynających naukę (Athanassopoulos i Shale 1997; Bessent i Bessent 1980; Chalos 1997; Johnes i Taylor 1990; Johnes 1996; Barrow 1991; Thanassoulis i Dunstan 1994), a w badaniach prowadzonych na poziomie indywidualnym przez średnią wyniku uzyskanego w procesie rekrutacji (zob. np. Johnes 2003; Portela i Thanassoulis 2001; Aitkin i Longford 1986; Yang i in. 1999; Bratti i in. 2004; Smith i Naylor 2001b). Miary związane z początkowymi osiągnięciami studentów oraz liczba studentów są często ze sobą łączone, by uwzględnić wpływ wywierany zarówno przez jakość studentów, jak i wielkość instytucji na produkowane wyniki (Athanassopoulos i Shale 1997).

Publikowanie danych finansowych dla instytucji edukacyjnych stwarza możliwość konstruowania miar nakładów, które obrazują nakład kapitału rzeczowego wykorzystywanego w procesie działalności edukacyjnej. W wielu badaniach jako miarę kapitału fizycznego stosuje się wydatki związane z różnymi nakładami (takie jak biblioteka czy sale informatyczne), przychód przeznaczony na różne cele, wartość aktywów czy poziom inwestycji (np. Harrison i Rouse 2002; Abbott i Doucouliagos 2003; Mante 2001; Mayston i Jesson 1988). Gdy dane te są dostępne, używa się również liczby książek czy komputerów jako alternatywy dla danych finansowych (Lovell i in. 1994; Ruggiero 1996, 2000).

Co więcej, poza już omówionymi nakładami pracy i kapitału ważnymi nakładami w procesie edukacji są również inne zmienne: pochodzenie społeczno-ekonomiczne, rasa, płeć, jak również zawód wykonywany przez rodziców oraz uzyskiwane przez nich dochody. Zmienne mające odzwierciedlać te cechy są tym samym często włączane na wszystkich poziomach analizy (zob. Bessent i Bessent 1980; Johnes i Taylor 1990 będące przykładami analiz prowadzonych na zagregowanym poziomie szkoły czy uniwersytetu; z kolei Goldstein (1997) oraz Bratti i in. (2004) przeprowadzili badania na poziomie indywidualnych – odpowiednio – szkół i uniwersytetów). Instytucje charakteryzują się dodatkowo własnym programem nauczania, zajęciami odbywającymi się poza lekcjami, lokalizacją, przepisami administracyjnymi i tymi dotyczącymi rekrutacji, jak również praktykami dydaktycznymi, które powinny być uznane za nakłady (Lovell i in. 1994; Goldstein i in. 1993; Goldstein i Thomas 1996; Yang i in. 2002; Johnes i Taylor 1990; Johnes 1996; Barrowe 1991).

Próba włączenia tak wielu miar nakładów i wyników w badania nad efektywnością instytucji powoduje kilka problemów. Po pierwsze, przy tak szerokim wachlarzu potencjalnych miar nakładów, z których wszystkie są *a priori* istotne, istnieje prawdopodobieństwo, że wystąpi między nimi problem współliniowości. Wiadomo, że problem ten pojawia się, gdy używa się metod statystycznych. Istnieją też pewne dowody na to, że może on wystąpić w metodach nieparametrycznych takich jak DEA, a jego konsekwencje różnią się w zależności od rozmiarów próby i poziomu analizy (Pedraja-Chaparro i in. 1999).

Konsekwencją wynikającą z trudności mierzenia nakładów jest to, że część z istotnych zmiennych obrazujących nakłady może zostać pominięta. Jako że w kontekście edukacji istnieje prawdopodobieństwo, że uwzględnione nakłady są skorelowane z tymi nieuwzględnionymi, pominięcie zmiennych będzie skutkowało w obciążonym wyniku w badaniach statystycznych (Bifulco i Bretschneider 2001). Pominięcie istotnych zmiennych nakładów może oddziaływać również na wartości wskaźników efektywności uzyskanych w ramach metody DEA.

Kolejny problem związany jest z wyborem, które zmienne obrazujące nakłady powinny zostać uwzględnione podczas konstruowania miar efektywności. Przyjęte

w ramach metod statystycznych praktyki polegają na użyciu wyłącznie tych nakładów, które są statystycznie istotne (choć może to być trudne do stwierdzenia, gdy występuje problem współliniowości). W przypadku metody DEA statystyczne testy istotności nie są możliwe do przeprowadzenia. Uwzględnienie wszystkich możliwych nakładów (istotnych i nieistotnych) byłoby poprawne tylko, jeśli na wyniki uzyskiwane metodą DEA nie oddziaływałaby liczba uwzględnionych nakładów. Tak się jednak nie dzieje (zob. np. Athanassopoulos i Shale 1997), a dodatkowe potencjalne problemy spowodowane przez współliniowość sprawiają, że strategia ta jest jeszcze mniej adekwatna. Jedno z rozwiązań polega na wybraniu zmiennych w oparciu o opinie eksperckie i wcześniejsze analizy statystyczne. Chalos i Cherian (1995) zaproponowali szereg kryteriów, które powinny zostać spełnione, by móc uwzględnić dany nakład w ramach metody DEA.

Po pierwsze, powinny istnieć empiryczne dowody przemawiające za włączeniem danej zmiennej. Po drugie, panel „ekspertki” powinien uznawać dany nakład za istotny. Po trzecie, analiza statystyczna powinna wykazywać istotny związek pomiędzy nakładami a określonymi wynikami. Wreszcie, nakłady nie mogą mieć wartości brakującej ani wynoszącej zero. Alternatywne i łatwiejsze w zastosowaniu podejście (Mancebon i Bandrés 1999; Mancebon i Mar Molinero 2000) polega na zaadoptowaniu strategii, z której wyprowadza się test statystyczny, analogiczny do testu F dla podzbioru zmiennych w regresji wielokrotnej. Po pierwsze, wszystkie możliwe nakłady i wyniki włączone w DEA muszą zostać wyszczególnione. Następnie wybiera się jedną zmienną (nakład lub wynik) i oblicza się efektywność DEA zarówno z uwzględnieniem tej zmiennej (model całkowity), jak i z jej wyłączeniem (model ograniczony). Trzeci krok polega na obliczeniu stosunku (oznaczonego przez ρ) efektywności modelu ograniczonego do efektywności modelu całkowitego dla wszystkich jednostek decyzyjnych. Na końcu, jeśli ρ różni się znacząco od 1 w dużej ilości jednostek, zmienna powinna zostać uwzględniona w metodzie DEA. Mancebon i Mar Molinero przyjęli zasadę, która mówi o tym, że: jeżeli mniej niż 15% jednostek decyzyjnych ma wartość $\rho \leq 0.9$, to dana zmienna może być usunięta z DEA.

Dalsze wątpliwości dotyczące doboru nakładów wynikają z rozróżnienia na nakłady, które mogą być pod kontrolą instytucji, i te, które się pod nią nie znajdują (np. czynniki środowiskowe). Istnieją dwa przeciwstawne podejścia do tego problemu. Pierwsze polega na uwzględnieniu w analizowaniu efektywności wszystkich nakładów niezależnie od tego, czy jednostka ma wpływ na ich kontrolowanie, czy nie (Cubbin i Tzanidakis 1998; Grosskopf 1996). Problem z takim podejściem polega na tym, że trudno jest uwzględnić fakt, że niektóre instytucje działają w niesprzyjających warunkach, tym samym instytucje te mogą zostać uznane za nieefektywne, gdy w rzeczywistości ich działanie jest konsekwencją otoczenia, w którym się znajdują. Dlatego w modelach, w których wszystkie nakłady są uwzględnione od początku,

nieefektywność może być przeszacowana¹⁸. Drugie podejście polega na zastosowaniu dwuetapowej procedury, w ramach której uzyskuje się wskaźniki efektywności dla danego zbioru instytucji przy użyciu odpowiednich metod (zob. część 4), a następnie efektywność ta jest w drugim etapie analizowana w odniesieniu do niekontrolowanych nakładów przy pomocy odpowiednich metod statystycznych¹⁹.

W wielu analizach efektywności przejęto drugie podejście (Bratti 2002; Bradley i in. 2001; Duncombe i in. 1997; Kirjavainen i Loikkanen 1998; Lovell i in. 1994; Mancebon i Mar Moliero 2000; McCarty i Yaisawarng 1993; Ramanathan 2001; Ray 1991; Grosskopf i Moutray 2001), jednak wciąż istnieje szereg zastrzeżeń, o których trzeba pamiętać w kontekście stosowania procedury dwuetapowej. Po pierwsze, metoda ta wymaga określenia funkcji opisującej zależności między niekontrolowanymi nakładami a funkcją produkcji (Ruggiero 1996). Tym samym błędy związane z błędną specyfikacją mogą pojawić się w drugim etapie.

Po drugie, wyniki efektywności mają zakres od 0 do 1, a więc albo powinny być odpowiednio przekształcone, albo na drugim etapie powinno użyć się odpowiednich metod, które rozwiązują problem tego, że zmienna zależna jest ograniczona np. analiza tobitowa (Lovell 1993).

Po trzecie, rozróżnienie, które z nakładów powinny zostać uwzględnione na pierwszym, a które na drugim etapie może być trudne. Lovell (1993) sugeruje, że zmienne uwzględniane w drugim etapie to wyłącznie te, nad którymi instytucja nie ma kontroli. Jednak definicja niekontrolowanej zmiennej może różnić się ze względu na kontekst analizy, dlatego decyzja odnośnie tego, które zmienne mają zostać użyte tylko w drugim etapie, musi uwzględniać ten aspekt.

Wreszcie, decyzja o zastosowaniu jedno- czy dwuetapowej procedury zależy od założenia odnośnie określonego stosunku między niekontrolowanymi nakładami a procesem produkcji i efektywnością. Dwuetapowa procedura zakłada, że zmienne (zazwyczaj niekontrolowane) nakładów w drugim etapie mają wpływ na efektywność przekształcania nakładów w wyniki. W przeciwieństwie do tego założenia procedura jednoetapowa zakłada, że wszystkie nakłady, włączając w to nakłady niekontrolowane, oddziałują na proces produkcji wyników z nakładów (ibid.).

¹⁸ Problem ten rozwiązał Ruggiero (1996), opracowawszy podejście w kontekście metody DEA, w którym każda jednostka decyzyjna posiada zbiór jednostek referencyjnych funkcjonujących w podobnych warunkach środowiskowych.

¹⁹ Równania dla dwóch etapów (pierwszy etap to wyznaczenie wskaźników efektywności, w drugim etapie dotyczy to determinant efektywności) mogą być szacowane osobno lub łącznie. Przykłady pierwszego z podejść są wymienione w tekście. Drugie podejście zostało niedawno rozwinięte i obejmuje zastosowanie metody opracowanej przez Battese i Coelli (1995), w której wykorzystuje się metodę największej wiarygodności do jednoczesnego oszacowania efektywności w pierwszym etapie według SFA i równania z drugiego etapu, gdzie uprzednio wyznaczona efektywność jest powiązana ze zmiennymi, które mogą wyjaśniać różnice wewnątrzinstytucjonalne. Przykład zastosowania tej metody w kontekście kosztów szkolnictwa wyższego w Wielkiej Brytanii przedstawiono w Stevens (2001).

Badacze powinni być zatem świadomi tego rozróżnienia i wybierać tę metodę, która odpowiada ich założeniom. W rzeczywistości o wyborze metody często decydują względy praktyczne. Na przykład, gdy występuje współliniowość między kontrolowanymi a niekontrolowanymi nakładami, jednoetapowa procedura jest bardziej adekwatna, jeśli w pierwszym kroku używa się metod statystycznych.

Istnieje niewiele porównań rezultatów otrzymanych w wyniku zastosowania jedno- i dwuetapowej procedury. Wyjątkiem jest badanie przeprowadzone przez McCarty'ego i Yaisawarnga (1993), którzy zaadaptowali dwuetapowe podejście, by przy użyciu modelu tobitowego w drugim etapie powiązać wskaźniki efektywności DEA ze zmiennymi, które są poza kontrolą jednostek decyzyjnych. Obliczyli w tym kontekście nowy zbiór efektywności w oparciu o drugi etap i pokazali, że wskaźniki efektywności uzyskane za pomocą podejścia dwuetapowego w niewielkim stopniu różnią się od tych uzyskanych za pomocą metody jednoetapowej, w której od początku w zbiorze nakładów były zarówno zmienne kontrolowane, jak i niekontrolowane.

Czasami głównym celem analizy jest ocena wpływu zmiennych środowiskowych, takich jak lokalizacja instytucji czy skutki zmian polityki wobec sektora, na wskaźniki efektywności. Jeśli używa się metod statystycznych, można włączyć odpowiednie zmienne zero-jedynkowe w celu ocenienia tych efektów. Natomiast jeśli używa się metody DEA, to nie jest do końca jasne, w jaki sposób zbadana być może siła tych zależności. Soterou i in. (1998) przeprowadzili badania na temat wpływu lokalizacji szkoły (region miejski czy wiejski) na efektywność jej działalności. W pierwszym kroku oszacowano wskaźniki efektywności DEA dla dwóch grup szkół podzielonych według lokacji. Następnie nakłady wszystkich nieefektywnych jednostek zostały dostosowane do postaci, którą przyjąłby, znajdując się na granicy produkcji charakterystycznej dla swojej lokalizacji, podczas gdy dla efektywnych jednostek nie wprowadzono żadnych zmian. Dla tak zmodyfikowanych danych przeprowadzono ponowne obliczenia wskaźników efektywności metodą DEA. Na końcu porównano wskaźniki efektywności uzyskane za pomocą DEA dla każdej z grup szkół i oceniono istotność ich różnic za pomocą testów nieparametrycznych Manna-Whitneya. Posługując się tą metodą, Soteriou i in. (1998) nie stwierdzili istotnych różnic w efektywności szkół wynikających z ich odmiennych lokalizacji. Podobna metoda użyta została do określenia wpływu programów wprowadzanych do szkół mających na celu podniesienie efektywności kształcenia (Diamond i Madewitz 1990).

Z pewnością zidentyfikowanie nakładów i wyników w kontekście edukacji rodzi poważne problemy. W dalszej części niniejszego tekstu przedstawimy wyniki badań empirycznych, pokazując, jaki wpływ na wyniki w zakresie wskaźników efektywności mają różne specyfikacje modeli oraz wybór metody analizy.

Wiarygodność wyników analiz efektywności

Sukces prowadzonych analiz efektywności zależy od dostarczenia wiarygodnych i spójnych rankingów instytucji, które nie są wrażliwe na specyfikę nakładów i wyników, ilość włączonych obserwacji, założenia co do efektu skali czy metody użyte do obliczenia wskaźników efektywności. Jest to niezmiernie istotne dla wiarygodności analiz i stało się przedmiotem badań. Niektóre z ich wyników opiszemy poniżej.

Badania na poziomie indywidualnym wykorzystujące metody statystyczne (takie jak modele wielopoziomowe czy modele probitowe) dowiodły, że większość instytucji nie może być wyodrębniana na gruncie wartości wskaźników efektywności. W badaniach z wykorzystaniem danych dla uczniów ze szkół przedziały ufności dla efektów wytwarzanych przez szkołę są bardzo szerokie, niewiele szkół może być zatem wyodrębnionych w oparciu o ich działanie (Goldstein i in. 1993; Goldstein i Spiegelhalter 1996; Goldstein i Thomas 1996), a znacząca różnica występuje jedynie między szkołami najmniej i najbardziej produktywnymi. W badaniach dotyczących nieukończenia studiów przez studentów efekty wytwarzane przez uniwersytet są znacząco różne od mediany tylko dla 6 najniżej i 12 najwyżej sklasyfikowanych uniwersytetów (Smith i Naylor 2001a). Podobne wyniki otrzymano, gdy wyniki działalności uniwersytetu były mierzone za pomocą wskaźników dla rynku pracy (Smith i in. 2000). Na podstawie tego można stwierdzić, że tworzenie rankingów instytucji w oparciu o efekty ich działalności, które nie różnią się pomiędzy poszczególnymi jednostkami, nie jest poprawne (Yang i in. 1999).

Przeprowadzono liczne badania, w których sprawdzono jak specyfikacja nakładów i wyników wpływa na wartość wskaźników efektywności. Badania te były prowadzone głównie z użyciem metody DEA w odróżnieniu od zastosowania metod statystycznych, dla których istotność zbioru nakładów i wyników może być testowana. W kontekście szkolnictwa wyższego wnioski obejmują zarówno twierdzenie, że uzyskane rankingi są w wystarczającym stopniu stabilne bez względu na specyfikę nakładów i/lub wyników (Tomkins i Green 1988; Abbot i Doucouliagos 2003; Johnes 2003), jak również, że na wyniki oddziałuje specyfikacja nakładów i wyników (Ahn i Seiford 1993; Johnes i Johnes 1992). W badaniach efektywności działalności naukowej prowadzonej na wydziałach ekonomii na brytyjskich uniwersytetach porównano wyniki otrzymane przy użyciu 192 różnych specyfikacji nakładów i wyników i na podstawie analizy skupień wykazano istnienie zaledwie dwóch odrębnych grup rezultatów (Johnes i Johnes 1993). Głównym czynnikiem różniącym obie grupy, w których uzyskano odmienne wyniki, było włączenie (lub brak takiego włączenia) do zbioru nakładów grantów badawczych *per capita*.

W badaniach funkcjonowania Training and Enterprise Councils (TECs) w Wielkiej Brytanii włączenie kolejnych zmiennych okazało się zwiększać liczbę TECs uznanych

za efektywne (Cubbin i Zamani 1996), pomimo że wyniki okazały się nie być wrażliwe na orientację (na nakłady lub na wyniki) modelu DEA.

W przypadku szkół średnich wyniki są również niejednoznaczne (Färe i in. 1989; Harrison i Rouse 2002; Smith i Mayston 1987; Kirjavainen i Loikkanen 1998; Mante 2001; Engert 1996). Z badania fińskich szkół średnich wyłania się interesujące spostrzeżenie, a mianowicie szkoły średnie zajmujące najwyższe i najniższe pozycje (pod względem wskaźników efektywności) są mniej podatne na zmianę specyfikacji modelu niż te plasujące się w środku stawki (Kirjavainen i Loikkanen 1998). Nie jest to zaskakujące, jeśli pod uwagę weźmiemy, że wyniki analizy przeprowadzonej na poziomie indywidualnym nie wykazały rozróżnienia pod względem efektywności między instytucjami zajmującymi miejsca w połowie stawki.

Te niejednoznaczne wnioski dotyczące wpływu specyfikacji modelu na ostateczne wyniki mogą być konsekwencją tego, co było celem przeprowadzonych badań: czy było nim zmierzenie tego samego typu efektywności w ramach różnych modeli (zob. np. Johnes 2003, w kwestii analizy efektywności wydziałów ekonomii pod kątem produkcji najwyższych stopni przy pomocy różnego zbioru nakładów i wyników), czy też mierzenie różnego rodzaju efektywności (dla przykładu, dwa modele wyszczególnione przez Athanassopoulou i Shale mierzą odpowiednio efektywność kosztową i efektywność wyników). Lovell i in. (1994) badają efektywność szkół w osiąganiu krótko-, średnio- i długoterminowych celów. Wskaźniki efektywności wyprowadzone z modelu dla celów krótkookresowych są w znacznej mierze różne od tych uzyskanych odpowiednio dzięki modelowi średnio- i długookresowemu. Co więcej, czynniki determinujące każdy z tych zbiorów wskaźników efektywności również się między sobą różnią. Podobna różnica w uzyskiwanych wskaźnikach efektywności zachodzi, gdy rozważa się różne typy efektywności. Wyniki pomiaru efektywności technicznej i efektywności wartościowej ukazują, że liczba jednostek efektywnych w drugim przywołanym sensie jest mniejsza niż liczba jednostek efektywnych technicznie (Korhonen i in. 2001).

Wyniki badań statystycznych również potwierdzają hipotezę, że wskaźniki efektywności różnią się, gdy wyprowadzane są na podstawie ewaluacji efektywności instytucji w osiąganiu różnych celów. W badaniu, w którym wzięto pod uwagę cztery różne miary działania uniwersytetu, wykorzystując kolejno dane na temat losów ekonomicznych absolwentów, wyników na dyplomach, odsetka osób nieukończających edukacji (zauważmy, że wskaźniki nieukończenia edukacji przedstawiane są w odwrotnej formie, tak by umożliwić spójną interpretację pomiędzy wskaźnikami wydajności i wskaźnika badań naukowych, współczynnik korelacji rang wynosił od -0.40 do 0.13²⁰,

²⁰ W rzeczywistości wskaźnik efektywności oparty na sukcesie absolwentów na rynku pracy jest negatywnie skorelowany z pozostałymi trzema wskaźnikami, z r od -0.07 do -0.40. Pozostałe wskaźniki efektywności są ze sobą skorelowane pozytywnie.

wskazując tym samym na duże zróżnicowanie uzyskanych wskaźników efektywności (Johnes 1996). Podobnie w badaniach przeprowadzonych na poziomie indywidualnym dzieci w wieku szkolnym wskaźniki efektywności otrzymane na podstawie wyników testów z czytania i wyników testów z matematyki są ze sobą skorelowane ($r = 0.62$), ale stopień tej korelacji nie jest wysoki (Sammons i in. 1993). Co więcej, w badaniu efektywności szkół publicznych w Nowym Jorku, które wykorzystywało stochastyczną metodę graniczną, średnia efektywność była uzależniona od definicji wyniku (zdefiniowano pięć miar wyników i średnia efektywność była w zakresie od 0.55 do 0.91). Ponadto pokazano, że różnice w przyjętych miarach nakładów odgrywają rolę w każdym z modeli (Kang i Greene 2002).

Nie ma dużo dowodów dotyczących wpływu pominięcia instytucji na stabilności funkcji produkcji, a co za tym idzie, wpływu na wskaźniki efektywności. Smith i Mayston (1987) pokazali zresztą, że nie jest to zaskakujące – wrażliwość wyników na pominięcie jednostek decyzyjnych zależy od tego, czy pominięta jednostka jest lub nie jest obserwacją odstającą. Jeżeli jest obserwacją odstającą, to jej pominięcie wpływa na wyniki. Jedną z metod oceny wpływu obserwacji jest posłużenie się metodą *jackknife*, w której dla określonego modelu DEA po kolei pomija się jedną jednostkę decyzyjną, a średni wskaźnik efektywności uzyskany z $(n-1)$ obliczeń porównuje się z wynikiem uzyskanym dla wszystkich n -jednostek. W badaniach rejonów szkolnych w Missouri i szkół średnich z Finlandii wyniki uzyskane za pomocą metody *jackknife* wskazują, że wskaźniki efektywności są stabilne (Färe i in. 1989; Bonesrønning i Rattsø 1994).

Badania na temat wpływu przyjętych korzyści skali na pomiar efektywności nie są liczne. W kontekście DEA Cubbin i Zamani (1996) zauważyli, że ilość efektywnych jednostek decyzyjnych zwiększa się (zgodnie z przewidywaniami) wraz z przyjęciem założenia o VSR niż CRS. Ahn i Seiford (1993) uważają, że otrzymane wyniki są niewrażliwe na założenie co do korzyści skali, ale są uzależnione od tego, czy użyje się modelu addytywnego, czy modelu multiplikatywnego. Athanassopoulos i Shale (1997) zauważyli, że wyniki dotyczące założeń korzyści skali różnią się w zależności od modelu: wyniki dla efektywności kosztowej mierzonej metodą DEA różnią się w zależności od tego, czy użyje się modelu CRS, czy modelu VRS; z kolei istnieje niewielki wpływ wyboru co do założeń korzyści skali na wskaźniki efektywności uzyskane metodą DEA zorientowaną na wyniki.

Badania wskazują, że wybór metody analizy ma wpływ na wartość wskaźników efektywności i rankingów instytucji, z których znaczna część dotyczy badań porównujących uzyskane wyniki za pomocą metody DEA i metod statystycznych (metody deterministyczne i stochastyczne). Banker i in. (1993) posługując się danymi pochodzącymi z symulacji w celu porównania COLS (z błędem deterministycznym) i DEA, odkryli, że skuteczność jednej metody w porównaniu z drugą zależy przede

wszystkim od rzeczywistego rozkładu wskaźników nieefektywności i wielkości próby. Wydaje się, że ta metoda jest bardziej skuteczna niż COLS dla małych prób ($n=25$). Jednak w sytuacji, w której pojawiają się istotne błędy pomiaru, żadna z metod nie działa szczególnie dobrze. Wyniki te potwierdzają Bifulco i Bretschneider (2001), którzy również wykorzystali symulowane dane. Pokazali, że dla złożonych zestawów danych, które charakteryzują się endogenicznością i zawierają błędy pomiaru (cechy, które w szczególności mogą się pojawić w zbiorach danych dotyczących edukacji) ani COLS, ani DEA nie są w stanie poprawnie wyznaczyć wskaźników dla więcej niż 31% szkół tak, żeby odpowiadały one wartościom rzeczywistym.

Szereg badań empirycznych został poświęcony porównaniom wyników otrzymanych przy użyciu MNK (z błędem deterministycznym) i metody DEA. W badaniach odnoszących się do kształcenia ekonomii w amerykańskich szkołach stwierdzono wysoką korelację pomiędzy wskaźnikami efektywności otrzymanymi dzięki trzem różnym modelom MNK i wskaźnikom efektywności otrzymanym przy użyciu DEA (r jest w zakresie od 0.712 do 0.822) (Diamond i Medewitz 1990). Bates (1993) stwierdza, że wskaźniki efektywności otrzymane dzięki MNK i DEA dla LEA w Wielkiej Brytanii są szczególnie podobne na szczycie i na dole rankingów efektywności. Jednak inne badanie dotyczące LEA w Wielkiej Brytanii wskazuje na raczej niski stopień korelacji (r jest w zakresie od 0.32 do 0.53) pomiędzy wskaźnikami efektywności uzyskanymi za pomocą dwóch modeli MNK i tymi otrzymanymi dzięki metodzie DEA. Dalsze dowody na to, że wybór metody (MNK czy DEA) oddziałuje na ranking efektywności, możemy znaleźć u Cubbina i Zamaniego (1996), którzy badali TEC.

DEA była również porównywana do stochastycznej metody granicznej. Gong i Sickles (1992), posługując się symulowanymi danymi, wskazują, że im bardziej niepoprawne określenie specyfikacji funkcji w stochastycznej metodzie granicznej i im wyższy stopień korelacji pomiędzy wskaźnikami nieefektywnością i zmiennymi objaśniającymi, tym bardziej atrakcyjna staje się metoda DEA. Badania empiryczne dostarczają niejednoznacznych wyników na spójność szacowanej efektywności w obu modelach, która waha się od niskiego stopnia korelacji w kontekście TEC w Wielkiej Brytanii (Cubbin i Zamani 1996) i brytyjskich uniwersytetów (Johnes 1998a) po stosunkowo wysoki stopień korelacji w kontekście LEA w Wielkiej Brytanii (Bates 1997). Chakraborty i in. (2001) pokazali, że rankingi dzielnic, w których zlokalizowane są szkoły wynikające z reszt uzyskanych z modelu tobitowego zastosowanego na drugim etapie analizy do powiązania wskaźników efektywności DEA z niekontrolowanymi nakładami, są podobne do tego, który został otrzymany przy użyciu stochastycznej metody granicznej z półnormalnym rozkładem błędów.

Metoda DEA zastosowana na poziomie indywidualnym została niedawno porównana z modelowaniem wielopoziomowym. Wstępna analiza wyników ujawnia wysoki stopień korelacji pomiędzy efektywnością uczniów w ramach jednej szkoły

oszacowanej odpowiednio przy użyciu metody DEA na poziomie indywidualnym i wielopoziomowego modelowania (Thanassoulis i in. 2003). Podobnie wysoki stopień korelacji występuje pomiędzy efektywnością uczniów w ramach wszystkich szkół oszacowaną na podstawie obu metod. Istnieje jednak potrzeba określenia zależności pomiędzy efektywnością instytucji uzyskaną dzięki wielopoziomowemu modelowaniu i DEA na poziomie indywidualnym.

Podobne niejednoznaczne wyniki zostały uzyskane przez porównanie wyników uzyskanych za pomocą metod statystycznych. Gdy porównuje się różne funkcje oszacowane przy pomocy MNK, podobieństwa były znacznie większe, gdy badano efektywność zajęć z ekonomii w Stanach Zjednoczonych (r od 0.31 do 0.956) niż w kontekście LEA w Wielkiej Brytanii (r od 0.21 do 0.60). Niemniej MNK i stochastyczna metoda graniczna wydają się dostarczać podobnych rankingów efektywności w kontekście LEA (Barrow 1991) i TEC (Cubbin i Zamani 1996), chociaż pierwsze z wymienionych badań (zgodnie z oczekiwaniami) wskazuje, że rzeczywiste wskaźniki efektywności są znacząco wyższe, gdy obliczane są za pomocą COLS, a nie modelu stochastycznego.

6. Zakończenie

Sposoby pomiaru efektywności w edukacji znacznie się rozwinęły od czasów, gdy brano pod uwagę tylko jeden wynik działalności instytucji. Przyjmuje się obecnie, że instytucje edukacyjne wytwarzają szereg wyników przy użyciu wielu nakładów, dlatego musi to zostać uwzględnione w tworzonych rankingach instytucji, które mają zobrazować ich efektywność.

Zaprezentowany w niniejszym rozdziale przegląd literatury jasno pokazuje, że każda osoba zainteresowana mierzaniem efektywności instytucji edukacyjnych ma przed sobą do wyboru cały wachlarz metod analizy i różnych sposobów podejścia do tematu. Ponadto otoczenie produkcyjne w sektorze edukacji różni się od tego, które spotykamy w bardziej tradycyjnych sektorach produkcji, co niewątpliwie sprawia problem przy próbach analizy efektywności. Przegląd empirycznych badań, w których zastosowano metody pomiaru efektywności w kontekście edukacji, wskazuje, że wyniki studiów nad efektywnością w postaci rankingów ocenianych instytucji mogą się różnić w zależności od doboru metod (na przykład: parametryczna i nieparametryczna); specyfikacji i pomiaru nakładów i wyników (jak również sposobu uwzględnienia nakładów, które nie znajdują się pod bezpośrednią kontrolą analizowanej instytucji); poziomu używanych danych (indywidualna jednostka, dyscyplina czy instytucja); założeń przyjętego modelu (postać funkcji produkcji lub kwestia tego, czy przyjmuje się istnienie korzyści skali produkcji). Ważne jest, aby pamiętać o tych kwestiach, analizując i interpretując jakiegokolwiek wyniki badań nad efektywnością instytucji.

Mocną stroną badań nad efektywnością stanowią nie tyle dostarczane przez nie „rankingi”, ale raczej bogaty zasób dostarczanych informacji. Informacje te zależą od zastosowanych metod. Metoda statystyczna może wskazywać na istotne czynniki warunkujące uzyskane rezultaty działalności edukacyjnej; metoda nieparametryczna może dostarczać informacji na temat realnych celów, jakie są do osiągnięcia przez nieefektywną instytucję, jak również informacji o podobnych (pod względem nakładów i wyników), ale lepiej funkcjonujących instytucjach, których praktyki nieefektywna organizacja może starać się naśladować.

Metody mierzenia efektywności są wciąż rozwijane i poszerzane, co z kolei gwarantuje zapotrzebowanie na ciągle badania empiryczne w zakresie ich zastosowań w kontekście edukacji.

Tłumaczenie: Joanna Wolszczak-Derlacz i Jakub Krzeski

Literatura

- Abbott, M. i Doucouliagos, C. (2003). The efficiency of Australian universities: A data envelopment analysis. *Economics of Education Review*. 22(1): 89–97.
- Ahn, T. i Seiford, L.M. (1993). Sensitivity of data envelopment analysis to models and variable sets in a hypothesis test setting: the efficiency of university operations (191–208). W: Y. Ijiri (red.), *Creative and Innovative Approaches to the Science of Management*. Westport, CT: Quorum Books.
- Ahn, T., Arnold, V., Charnes, A. i Cooper, W.W. (1989). DEA and ratio efficiency analyses for public institutions of higher learning in Texas. *Research in Governmental and Nonprofit Accounting*. 5: 165–185.
- Aigner, D.J., Lovell, C.A.K. i Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*. 6(1): 21–37.
- Aitkin, M. i Longford, N. (1986). Statistical modelling in school effectiveness studies. *Journal of the Royal Statistical Society, Series A*. 149: 1–43.
- Arcelus, F.J. i Coleman, D.F. (1997). An efficiency review of university departments. *International Journal of Systems Science*. 28(7): 721–729.
- Athanassopoulos, A.D. i Shale, E. (1997). Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis. *Education Economics*. 5(2): 117–34.
- Avkiran, N.K. (2001). Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences*. 35(1): 57–80.
- Banker, R.D., Charnes, A. i Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*. 30(9): 1078–1092.

- Banker, R.D., Charnes, A., Cooper, W.W., Swarts, J. i Thomas, D. (1989). An introduction to data envelopment analysis with some of its models and their uses. *Research in Governmental and Nonprofit Accounting*. 5: 125–163.
- Banker, R.D., Gadh, V.M. i Gorr, W.L. (1993). A Monte Carlo comparison of two production frontier estimation techniques: corrected ordinary least squares and data envelopment analysis. *European Journal of Operational Research*. 67: 332–343.
- Barrow, M. (1991). Measuring the Local Education Authority performance: A frontier approach. *Economics of Education Review*. 10(1): 19–27.
- Bates, J.M. (1993). The efficiency of Local Education Authorities. *Oxford Review of Education*. 19(3): 277–289.
- Bates, J.M. (1997). Measuring predetermined socio-economic “inputs” when assessing the efficiency of educational outputs. *Applied Economics*. 29(1): 85–93.
- Battese, G.E. i Coelli, T.J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*. 20: 325–332.
- Beasley, J.E. (1990). Comparing university departments. *Omega*. 18: 171–183.
- Beasley, J.E. (1995). Determining the teaching and research efficiencies’, *Journal of the Operational Research Society*. 46(4): 441–452.
- Belfield, C.R. i Fielding, A. (2001). Measuring the relationship between resources and outcomes in higher education in the UK. *Economics of Education Review*. 20: 589–602.
- Bessent, A.M. i Bessent, E.W. (1980). Determining the comparative efficiency of schools through data envelopment analysis. *Educational Administration Quarterly*. 16(2): 57–75.
- Bessent, A.M., Bessent, E.W., Kennington, E.W. i Reagan, B. (1982). An application of mathematical programming to assess the productivity in the Houston independent school district. *Management Science*. 28: 1355–1367.
- Bifulco, R. i Bretschneider, S. (2001). Estimating school efficiency: A comparison of methods using simulated data. *Economics of Education Review*. 20: 417–429.
- Bonesrønning, H. i Rattsø, J. (1994). Efficiency variation among the Norwegian high schools: Consequences of equalization policy. *Economics of Education Review*. 13(4): 289–304.
- Bradley, S., Johnes, G. i Millington, J. (2001). The effect of competition on the efficiency of secondary schools in England. *European Journal of Operational Research*. 135: 545–568.
- Bratti, M. (2002). Does the choice of university matter? A study of the differences across UK universities in life sciences students’ degree performance. *Economics of Education Review*. 21: 431–443.
- Bratti, M., McKnight, A., Naylor, R. i Smith, J. (2004). Higher education outcomes, graduate employment and university performance indicators. *Journal of the Royal Statistical Society: Series A*. 167: 475–496.
- Breu, T.M. i Raab, R.L. (1994). Efficiency and perceived quality of the nation’s ‘Top 25’ national universities and national liberal arts colleges: An application of data envelopment analysis to higher education. *Socio-Economic Planning Sciences*. 28(1): 33–45.

- Chakraborty, K., Biswas, B. i Lewis, W.C. (2001). Measurement of technical efficiency in public education: A stochastic and nonstochastic production frontier approach. *Southern Economics Journal*. 67(4): 889–905.
- Chalos, P. (1997). An examination of budgetary inefficiency in education using data envelopment analysis. *Financial Accountability and Management*. 13(1): 55–69.
- Chalos, P. i Cherian, J. (1995). An application of data envelopment analysis to public sector performance measurement and accountability. *Journal of Accounting and Public Policy*. 14: 143–160.
- Charnes, A., Cooper, W.W. i Rhodes, E. (1978). Measuring the efficiency of DMUs. *European Journal of Operational Research*. 2: 429–444.
- Charnes, A., Cooper, W.W. i Rhodes, E. (1979). Measuring the efficiency of decision making units: A short communication. *European Journal of Operational Research*. 3(4): 339
- Chizmar, J.F. i Zak, T.A. (1983). Modeling multiple outputs in educational production functions. *American Economic Review, Papers and Proceedings*. 73(2): 18–22.
- Coelli, T., Rao, D.S.P. i Battese, G.E. (1998). *An Introduction to Efficiency and Productivity Analysis*. Norwell, MA: Kluwer Academic.
- Cubbin, J. i Tzanidakis, G. (1998). Regression versus data envelopment analysis for efficiency measurement: An application to the England and Wales regulated water industry. *Utilities Policy*. 7: 75–85.
- Cubbin, J. i Zamani, H. (1996). A comparison of performance indicators for Training and Enterprise Councils in the UK. *Annals of Public Choice and Cooperative Economics*. 67(4): 603–632.
- Dantzig, G.B. (1951). Maximization of a linear function of variables subject to linear inequalities. W: T.C. Koopmans (red.), *Activity Analysis of Production and Allocation*. New York: Wiley.
- Deller, S.C. i Rudnicki, E. (1993). Production efficiency in elementary education: The case of Maine public schools. *Economics of Education Review*. 12(1): 45–57.
- Diamond, A. i Medewitz, J.N. (1990). Use of data envelopment analysis in an evaluation of the efficiency of the DEEP program for economic education. *Journal of Economic Education*. 21: 337–354.
- Duncombe, W., Miner, J. i Ruggiero, J. (1997). Empirical evaluation and bureaucratic models of inefficiency. *Public Choice*. 93: 1–18.
- Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S. i Shale, E.A. (2001). Pitfalls and protocols in data envelopment analysis. *European Journal of Operational Research*. 132(2): 245–259.
- El-Mahgary, S. i Lahdelma, R. (1995). Data envelopment analysis: Visualizing the results. *European Journal of Operational Research*. 85: 700–710.
- Engert, F. (1996). The reporting of school district efficiency: The adequacy of ratio measures. *Public Budgeting and Financial Management*. 8(2): 247–271.

- Färe, R., Grosskopf, S. i Lovell, C.A.K. (1988). An indirect efficiency approach to the evaluation of producer performance. *Journal of Public Economics*. 37(1): 71–89.
- Färe, R., Grosskopf, S., Norris, M. i Zhang, Z. (1994). Productivity growth, technical progress and efficiency change in industrialized countries. *American Economic Review*. 84(1): 66–83.
- Färe, R., Grosskopf, S. i Weber, W.L. (1989). Measuring school district performance. *Public Finance Quarterly*. 17(4): 409–428.
- Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A*. 120: 253–281.
- Førsund, F.R., Lovell, C.A.K. i Schmidt, P. (1980). A survey of frontier production functions and of their relationship to efficiency measurement. *Journal of Econometrics*. 13: 5–25.
- Friedman, L. i Sinuany-Stern, Z. (1997). Scaling units via the canonical correlation analysis in the data envelopment analysis context. *European Journal of Operational Research*. 100: 629–637.
- Ganley, J.A. i Cubbin, J.S. (1992). *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*. Amsterdam: North-Holland.
- Geva-May, I. (2001). Higher education and attainment of policy goals: Interpretations for efficiency indicators in Israel. *Higher Education*. 42(3): 265–305.
- Goldstein, H. (1995). *Multilevel Statistical Models*. London: Wiley and Sons.
- Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement*. 8(4): 369–395.
- Goldstein, H. i Sammons, P. (1997). The influence of secondary and junior schools on sixteen year examination performance: A cross-classified multilevel analysis. *School Effectiveness and School Improvement*. 8(2): 219–230.
- Goldstein, H. i Spiegelhalter, D.J. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society, Series A*. 159(3): 385–443.
- Goldstein, H. i Thomas, S. (1996). Using examination results as indicators of school and college performance. *Journal of the Royal Statistical Society, Series A*. 159(1): 149–163.
- Goldstein, H., Rabash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D. i Thomas, S. (1993). A multilevel analysis of school examination results. *Oxford Review of Education*. 19(4): 425–433.
- Gong, B.-H. i Sickles, R.C. (1992). Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data. *Journal of Econometrics*. 51: 259–284.
- Gray, J., Jesson, D. i Jones, B. (1984). Predicting differences in examination results: Does school organisation matter? *Oxford Review of Education*. 10: 45–68.
- Grosskopf, S. (1996). Statistical inference and nonparametric efficiency: A selective survey', *Journal of Productivity Analysis*. 7: 161–176.

- Grosskopf, S. i Moutray, C. (2001). Evaluating performance in Chicago public high schools in the wake of decentralization. *Economics of Education Review*. 20: 1–14.
- Häkkinen, I., Kirjavainen, T. i Uusitalo, R. (2003). School resources and student achievement revisited: New evidence from panel data. *Economics of Education Review*. 22(3): 329–335.
- Haksever, C. i Muragishi, Y. (1998). Measuring value in MBA programmes. *Education Economics*. 6(1): 11–25.
- Halme, M., Joro, T., Korhonen, P., Salo, S. i Wallenius, J. (1999). A value efficiency approach to incorporating preference information in data envelopment analysis. *Management Science*. 45: 103–115.
- Harrison, J. i Rouse, P. (2002). Measuring the performance of Auckland secondary schools: A pilot study using data envelopment analysis. <http://orsnz.org.nz/conf37/Papers/Rouse.pdf> [30.06.2018].
- Hashimoto, K. i Cohn, E. (1997). Economies of scale and scope in Japanese private universities. *Education Economics*. 5(2): 107–115.
- Izadi, H., Johnes, G., Oskrochi, R. i Crouchley, R. (2002). Stochastic frontier estimation of a CES cost function: The case of higher education in Britain. *Economics of Education Review*. 21: 63–71.
- Jesson, D., Mayston, D. i Smith, P. (1987). Performance assessment in the education sector: Educational and economic perspectives. *Oxford Review of Education*. 13: 249–267.
- Johnes, G. (1996). Multi-product cost functions and the funding of tuition in UK universities. *Applied Economics Letters*. 3(9): 557–561.
- Johnes, G. (1997). Costs and industrial structure in contemporary British higher education. *Economic Journal*. 107: 727–737.
- Johnes, G. (1998a). The costs of multiproduct organizations and the heuristic evaluation of industrial structure. *Socio-Economic Planning Sciences*. 32(3): 199–209.
- Johnes, G. (1998b). Corrigendum: Costs and industrial structure in contemporary British higher education. *Economic Journal*. 108: 1275.
- Johnes, G. (1999). The management of universities: Scottish Economic Society/Royal Bank of Scotland Annual Lecture. *Scottish Journal of Political Economy*. 46: 505–522.
- Johnes, G. i Johnes, J. (1992). Apples and oranges: The aggregation problem in publications analysis. *Scientometrics*. 25(2): 353–365.
- Johnes, G. i Johnes, J. (1993). Measuring the research performance of UK economics departments: An application of data envelopment analysis. *Oxford Economic Papers*. 45: 332–347.
- Johnes, J. (1990). Unit costs: Some explanations of the differences between UK universities. *Applied Economics*. 22: 853–862.
- Johnes, J. (1996). Performance assessment in higher education in Britain. *European Journal of Operational Research*. 89: 18–33.

- Johnes, J. (2003). Measuring teaching efficiency in higher education: An application of data envelopment analysis to graduates from UK universities 1993. Discussion paper EC7/03. Department of Economics, Lancaster University.
- Johnes, J. i Taylor, J. (1987). Degree quality: An investigation into differences between universities. *Higher Education*. 16: 581–602.
- Johnes, J. i Taylor, J. (1989a). Undergraduate non-completion rates: Differences between UK universities. *Higher Education*. 18: 209–225.
- Johnes, J. i Taylor, J. (1989b). The first destination of new graduates: Comparisons between universities. *Applied Economics*. 21(3): 357–374.
- Johnes, J. i Taylor, J. (1990). *Performance Indicators in Higher Education: UK Universities*. Milton Keynes: Open University Press and The Society for Research into Higher Education.
- Johnes, J. i Taylor, J. (1992). The 1989 research selectivity exercise: A statistical analysis of differences in research rating between universities at the cost centre level. *Higher Education Quarterly*. 46(1): 67–87.
- Johnes, J., Taylor, J. i Francis, B. (1993). The research performance of UK universities: A statistical analysis of the results of the 1989 research selectivity exercise. *Journal of the Royal Statistical Society, Series A*. 156(2): 271–286.
- Jondrow, J., Lovell, C.A.K., Materou, I.S. i Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*. 19: 233–238.
- Jones, J.T. i Zimmer, R.W. (2001). Examining the impact of capital on academic achievement. *Economics of Education Review*. 20: 577–588.
- Kang, B.-G. i Greene, K.V. (2002). The effects of monitoring and competition on public education outputs: A stochastic frontier approach. *Public Finance Review*. 30(1): 3–26.
- Kirjavainen, T. i Loikkanen, H.A. (1998). Efficiency differences of Finnish senior secondary schools: An application of data envelopment analysis and Tobit analysis. *Economics of Education Review*. 17(4): 377–394.
- Korhonen, P., Tainio, R. i Wallenius, J. (2001). Value efficiency analysis of academic research. *European Journal of Operational Research*. 130: 121–132.
- Lovell, C.A.K. (1993). Production frontiers and productive efficiency (3–67). W: H.O. Fried, C.A.K. Lovell i S.S. Schmidt (red.), *The Measurement of Productive Efficiency*. Oxford: Oxford University Press.
- Lovell, C.A.K., Walters, L.C. i Wood, L.L. (1994). Stratified models of education production using modified data envelopment analysis and regression analysis. W: A. Charnes, W.W. Cooper, A.Y. Lewin i L.M. Seiford (red.), *Data Envelopment Analysis: Theory, Methodology and Applications*. Dordrecht: Kluwer Academic.
- Madden, G., Savage, S. i Kemp, S. (1997). Measuring public sector efficiency: A study of economics departments at Australian Universities. *Education Economics*. 5(2): 153–167.

- Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de Estadística*. 4: 209–242.
- Mancebon, M.J. i Bandrés, E. (1999). Efficiency evaluation in secondary schools: The key role of model specification and of *ex post* analysis of results. *Education Economics*. 7(2): 131–152.
- Mancebon, M.J. i Mar Molinero, C. (2000). Performance in primary schools. *Journal of the Operational Research Society*. 51(7): 843–854.
- Mante, B. (2001). Measuring the performance of state secondary schools in Victoria: An application of data envelopment analysis. *Education Research and Perspectives*. 28(1): 105–132.
- Mante, B. i O'Brien, G. (2002). Efficiency measurement of Australian public sector organisations: The case of state secondary schools in Victoria. *Journal of Educational Administration*. 40(3): 274–296.
- Mayston, D. i Jesson, D. (1988). Developing models of educational accountability. *Oxford Review of Education*. 14: 321–339.
- McCarty, T.A. i Yaisawarng, S. (1993). Technical efficiency in New Jersey school districts (271–287). W: H.O. Fried, C.A.K. Lovell i S.S. Schmidt (red.), *The Measurement of Productive Efficiency*. Oxford: Open University Press.
- Montmarquette, C., Mahseredjian, S. i Houle, R. (2001). The determinants of university dropouts: A bivariate probability model with sample selection. *Economics of Education Review*. 20(5): 475–484.
- Naylor, R., Smith, J. i McKnight, A. (2000). Occupational earnings of graduates: Evidence for the 1993 UK university population. Department of Economics, University of Warwick. https://www.researchgate.net/publication/228336591_Occupational_earnings_of_graduates_evidence_for_the_1993_UK_university_population [30.06.2018].
- Naylor, R., Smith, J. i McKnight, A. (2002). Why is there a graduate earnings premium for students from independent schools? *Bulletin of Economic Research*. 54(4): 315–340.
- Norman, M. i Stoker, B. (1991). *Data Envelopment Analysis: The Assessment of Performance*. Chichester: Wiley.
- Ondrich, J. i Ruggiero, J. (2001). Efficiency measurement in the stochastic frontier model. *European Journal of Operational Research*. 129: 434–442.
- Pedraja-Chaparro, F., Salinas-Jimenez, J. i Smith, P. (1999). On the quality of the data envelopment analysis model. *Journal of the Operational Research Society*. 50: 636–644.
- Peraita, C. i Sanchez, M. (1998). The effect of family background on children's level of schooling and attainment in Spain. *Applied Economics*. 30: 1327–1334.
- Portela, M.C.A.S. i Thanassoulis, E. (2001). Decomposing school and school-type efficiency. *European Journal of Operational Research*. 132(2): 357–373.
- Ramanathan, R. (2001). A data envelopment analysis of comparative performance of schools in the Netherlands. *Opsearch*. 38(2): 160–182.

- Ray, S.C. (1991). Resource use efficiency in public schools: A study of Connecticut data. *Management Science*. 37: 1620–1628.
- Ruggiero, J. (1996). On the measurement of technical efficiency in the public sector. *European Journal of Operational Research*. 90: 553–565.
- Ruggiero, J. (2000). Nonparametric estimation of returns to scale in the public sector with an application to the provision of educational services. *Journal of the Operational Research Society*. 51: 906–912.
- Ruggiero, J. i Vitaliano, D.F. (1999). Assessing the efficiency of public schools using DEA and frontier regression. *Contemporary Economic Policy*. 17(3): 321–331.
- Sammons, P., Nuttall, D. i Cuttance, D. (1993). Differential school effectiveness: Results from a reanalysis of the Inner London Education Authority's junior school project data. *British Educational Research Journal*. 19: 381–405.
- Sarrico, C.S. i Dyson, R.G. (2000). Using data envelopment analysis for planning in UK universities – an institutional perspective. *Journal of the Operational Research Society*. 51: 789–800.
- Sarrico, C.S., Hogan, S.M., Dyson, R.G. i Athanassopoulos, A.D. (1997). Data envelopment analysis and university selection. *Journal of the Operational Research Society*. 48: 1163–1177.
- Schmidt, P. (1985/1986). Frontier production functions. *Econometric Reviews*. 4(2): 289–328.
- Sengupta, J.K. (1987). Production frontier estimation to measure efficiency: A critical evaluation in light of data envelopment analysis. *Managerial and Decision Economics*. 8: 93–99.
- Sengupta, J.K. (1990). Tests of efficiency in DEA. *Computers Operations Research*. 17(2): 123–132.
- Sengupta, J.K. (1999). The measurement of dynamic productive efficiency. *Bulletin of Economic Research*. 51(2): 111–124.
- Sengupta, J.K. i Sfeir, R.E. (1986). Production frontier estimates of scale in public schools in California. *Economics of Education Review*. 5(3): 297–307.
- Sengupta, J.K. i Sfeir, R.E. (1988). Efficiency measurement by data envelopment analysis with econometric applications. *Applied Economics*. 20: 285–293.
- Sickles, R.C., Good, D.H. i Getachew, L. (2002). Specification of Distance Functions Using Semi- and Nonparametric Methods with an Application to the Dynamic Performance of Eastern and Western European Air Carriers. *Journal of Productivity Analysis*. 17(1–2): 133–155.
- Simar, L. i Wilson, P.W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*. 44(1): 49–61.
- Simar, L. i Wilson, P.W. (2004). Performance of the Bootstrap for DEA Estimators and Iterating the Principle (265–298). W: W.W. Cooper, L.M. Seiford i J. Zhu (red.), *Handbook on Data Envelopment Analysis*. New York: Springer.

- Smith, J. i Naylor, R. (2001a). Dropping out of university: A statistical analysis of the probability of withdrawal for UK university students. *Journal of the Royal Statistical Society, Series A*. 164(2): 389–405.
- Smith, J. i Naylor, R. (2001b). Determinants of degree performance in UK universities: A statistical analysis of the 1993 student cohort. *Oxford Bulletin of Economics and Statistics*. 63: 29–60.
- Smith, J., McKnight, A. i Naylor, R. (2000). Graduate employability: Policy and performance in higher education in the UK. *Economic Journal*. 110: F382–F411.
- Smith, P. i Mayston, D. (1987). Measuring efficiency in the public sector. *Omega*. 15: 181–189.
- Soteriou, A., Karahanna, E., Papanastasiou, C. i Diakourakis, M.S. (1998). Using data envelopment analysis to evaluate the efficiency of secondary schools: The case of Cyprus. *International Journal of Educational Management*. 12(2): 65–73.
- Stevens, P.A. (2001). The determinants of economic efficiency in English and Welsh universities. Discussion paper no. 185. National Institute of Economic and Social Research, London.
- Thanassoulis, E. (1996). Altering the bias in differential school effectiveness using data envelopment analysis. *Journal of the Operational Research Society*. 47: 882–894.
- Thanassoulis, E. (1999). Setting achievement targets for school children. *Education Economics*. 7(2): 101–119.
- Thanassoulis, E. i Dunstan, P. (1994). Guiding schools to improved performance using data envelopment analysis: An illustration with data from a local education authority. *Journal of the Operational Research Society*. 45(11): 1247–1262.
- Thanassoulis, E. i Portela, M.C.A.S. (2002). School outcomes: Sharing the responsibility between pupil and school. *Education Economics*. 10(2): 183–207.
- Thanassoulis, E., Simpson, G., Battisti, G. i Charlesworth-May, A. (2003). DEA and multi-level modelling as alternative methods for assessing pupil and school performance. Discussion Paper. Aston Business School.
- Tomkins, C. i Green, R. (1988). An experiment in the use of DEA for evaluating the efficiency of UK university departments of accounting. *Financial Accountability and Management*. 4(2): 147–164.
- Woodhouse, G. i Goldstein, H. (1988). Educational performance indicators and LEA league tables. *Oxford Review of Education*. 14: 301–319.
- Worthington, A. (2001). An empirical survey of frontier efficiency measurement techniques in education. *Education Economics*. 9(3): 245–268.
- Yang, M. i Woodhouse, G. (2001). Progress from GCSE to A and AS level: Institutional differences and gender differences, and trends over time. *British Educational Research Journal*. 27(3): 24–46.
- Yang, M., Goldstein, H., Browne, W. i Woodhouse, G. (2002). Multivariate multilevel analyses of examination results. *Journal of the Royal Statistical Society, Series A*. 165(1): 137–153.

Yang, M., Goldstein, H., Rath, T. i Hill, N. (1999). The use of assessment data for school improvement purposes. *Oxford Review of Education*. 25(4): 469–483.

Aneks. Wykaz badań empirycznych*

Efficiency measurement

ABSTRACT: This article aims to identify and present the most common methods of efficiency measurement in the context of assessing how well institutions of education and higher education perform and to review the empirical studies of efficiency measurement at all levels of education.

KEYWORDS: efficiency, higher education, education, measurement, nonparametric technique, DEA

CYTOWANIE: Johnes, J. (2018). Pomiar efektywności. *Nauka i Szkolnictwo Wyższe* 2(52): 17–81. DOI: 10.14746/nisw.2018.2.1

JILL JOHNES – profesor na Uniwersytecie w Huddersfield, Wielka Brytania gdzie jest dziekanem Huddersfield Business School, jest także profesorem wizytującym/honorowym na Uniwersytecie w Lancaster, na którym pracowała we wcześniejszych latach. Jej liczne prace badawcze koncentrują się przede wszystkim na ocenie efektywności funkcjonowania organizacji (szkół wyższych oraz banków). Jest autorem wielu cytowanych artykułów oraz współautorem wpływowych książek, w tym: *Performance Indicators in Higher Education* (1990) oraz *International Handbook on the Economics of Education* (2004). Dzięki swojej pracy naukowej zyskała reputację i prestiż w dziedzinie badań nad efektywnością edukacji. Współredagowała specjalne wydanie czasopisma *Journal of the Operational Research Society* (2016) poświęcone badaniom nad efektywnością edukacji. Współpracuje z wieloma organizacjami międzynarodowymi w zakresie badań nad efektywnością, była współorganizatorką konferencji: *Efficiency in Education*, Londyn 2014 oraz *Efficiency in Education, Health and other Public Services*, Huddersfield 2018.

E-mail: J.Johnes@hud.ac.uk

* Z powodów ograniczonej liczby stron czasopisma, odsyłamy zainteresowanych czytelników do oryginalnego tekstu.