Studiengang Mathematik - Methoden und Modelle

Präsentation zu Modul 1 (Prüfung zum Kurs 01212)

Prüfer/Betreuer: Prof. Dr. Hochstättler

Benchmarking im Dienstleistungssektor mit Linearer Programmierung (Vergleich zweier Methoden)

Tobias Weber, Matrikelnummer 6352367

Maria-Theresien-Str. 16

66954 Pirmasens

vorgelegt im Oktober 2005

Inhaltsverzeichnis

1 Einleitung

The mathematics and the metaphysics,
fall to them as you find your stomach serves you.
No profit grows where is no pleasure ta'en,
in brief, sir, study what you most affect.
William Shakespeare
The Taming of the Shrew , Act I. Scene I.

In der vorliegenden Arbeit soll das Themengebiet des Kurses 01212 Lineare Optimierung in einer praktischen Anwendung betrachtet werden. Die Aufgabenstellung, einen Methodenvergleich für Benchmarking im öffentlichen Sektor anzustellen, ist für das Format der Prüfungsleistung zu weitgehend, dementsprechend wurde das Thema in Absprache auf einen Vergleich der im Artikel **"Benchmarking und Leistungsvergleiche in Filialnetzen großer Unternehmen"**[1] verfolgten Ansätze eingeschränkt. Eine besonderere Beachtung soll dabei der Data Envelopment Analysis (DEA) zukommen.

Diese Arbeit ist wie folgt aufgebaut: Im Hauptteil werden die von BHHH verfolgten Modellierungen gegenübergestellt. Die beiden Methoden bestehen aus der "klassischen" Linearen Optimierung einerseits und der DEA andererseits. DEA als Methode wird in Verbindung mit den Abschnitten E bis G der Quelle näher vorgestellt. Der Hauptteil endet mit weiteren vergleichenden Aspekten des Benchmarkings. Im dritten Abschnitt wird eine kurze abschließende Bewertung gegeben.

[1] Der Artikel wurde in der Zeitschrift für Betriebswirtschaft veröffentlicht und ist im Literaturverzeichnis mit [ABHH96] angegeben.
Die Seitenangaben in dieser Arbeit beziehen sich auf den unter
http://www.zaik.uni-koeln.de/~paper/preprints.html?show=zpr95-197
frei verfügbaren Vorabdruck. Entsprechend der damaligen Initialien der Autoren Bachem, Hecht, Heesen und Hochstättler wird diese Quelle mit BHHH abgkürzt.

2 Benchmarking-Methoden bei BHHH

2.1 Vorbemerkung

Der zugrunde liegende Aufsatz beschreibt unterschiedliche Möglichkeiten Benchmarking und Leistungsvergleiche in Unternehmungen zu modellieren, die über mehrere Filialen verfügen, die unter Einsatz einer Inputgröße (Arbeitszeit) durch ihre Produktion eine Anzahl von Outputgrößen erzeugen. Ein Leistungsvergleich zielt dabei auf Bestimmung von Arbeitsproduktivitäten ab: in einer Filiale kennt man nur die eingesetzte Gesamtarbeitszeit und die Anzahl der damit erzielten Dienstleistungen und Produkte; eine Information über die in einer Filiale zur Erstellung einer bestimmten Outputmenge verwendeten Zeit fehlt hingegen. Kernfrage ist damit: Wie lassen sich die Ansätze der Linearen Programmierung nutzen, um Aussagen über den Gewichtungsvektor zu treffen, mit dem die Outputgrößen bewertet werden? Mit Kenntnis dieses Gewichtungsvektors werden dann Benchmarking-Aussagen möglich.

2.2 Einsatz der klassischen Optimierung

2.2.1 Basismodell

Zielsetzung des ersten Modells ist es, einen Gewichtungsvektor $x = (x_1, ..., x_n)^T$ zu finden, "bei dem die Abweichungen der Produktivitäten der Filialen untereinander möglichst klein sind"[2]. x_j bezeichnet die Gewichtung für Produkt j; es ist zu bemerken, dass diese noch zu findenden Gewichtungen für alle Filialen gleich sind.[3] Ausgehend von m Filialen, die n Produkte erzeugen, erhält man eine Matrix $A = (a_{ij})_{1 \leq i \leq m\,;1 \leq j \leq n}$ mit den Häufigkeiten von Produkt j in Filiale i. Bekannt ist die alleinige Inputgröße c_i, die den Mitarbeitereinsatz (z. B. in Arbeitsstunden) darstellt, für das gesamte Unternehmen ist dementsprechend $c = (c_1, ..., c_m)^t$. Für Filiale i erhält man unter Verwendung des Gewichungsvektors x die Arbeitsproduktivität $p_i = \frac{1}{c_i} \sum_{j=1}^{n} a_{ij} x_j$. Der unbekannte Gewichtungsvektor x lässt nun für jede Filiale einen Vergleich von $A_{i.}x$ und der aufgewendeten Arbeit c_i zu: Ist $A_{i.}x < c_i$, wird durch Hinzufügen einer nichtnegativen Schlupfvariablen r_i die Ungleichung in eine Gleichung überführt; ist $A_{i.}x > c_i$ wird analog ein s_i eingeführt. Man erhält für das Gesamtunternehmen folgenden Zusammenhang: $Ax + r - s = c$, und sucht mit mehreren Zwischenüberlegungen den "Vektor x, der die Summe der

[2] [BHHH95] S.4

[3] TW: Produktivitätsunterschiede werden damit nicht in den Fokus gerückt. Eine "Grundeffizienz" wird auf hohem Niveau unterstellt.

Absolutbeträge der Abweichungen von $Ax = c$ minimiert."[4] Das zugehörige LP erscheint schlicht:

$$\begin{aligned} min \quad & \mathbb{1}^T r \\ & Ax + r - s = c \\ & \mathbb{1}^T r = \mathbb{1}^T s \\ & x, r, s \geq 0. \end{aligned}$$

Durch Lösen des LP erhält man den Gewichtungvektor x und kann jeder Filiale, deren Produktivität nicht exakt 1 ist, ein positives s_i oder r_i zuordnen. Eine Benchmarking-Aussage erhält man durch Ausrechnen der Produktivitäten $p_i = \frac{A_{i.}x}{c_i}$. BHHH schlagen nun vor, entsprechend r und s die Mitarbeiter(stunden) innerhalb der Filialen umzuverteilen, bzw. zumindest die Mitarbeiter(stunden) in den Filialen mit positivem r_i entsprechend einzusparen. Stets basierend auf dem bezüglich des LP optimalen Gewichtungsvektors x lassen sich nicht nur Sollproduktivitäten für die einzelnen Filialen formulieren, sondern auch die durch die Umschichtungen erzielten Produktivitätszuwächse im Gesamtunternehmen quantifizieren.

2.2.2 Erweiterungen

Das bei BHHH unter C. beschriebene "Modifizierte Modell"[5] setzt das Basismodell ein, und führt einen vom Management festlegbaren Parameters "Mindestfreisetzungskapazität k" in das Basismodell ein. Da gegenüber dem Basismodell nicht das Benchmarking der Filialen im Fokus steht, sondern die bezüglich des erweiterten LP optimale Umverteilung der Mitarbeiterkapazitäten, wird hier auf das modifizierte Modell nicht näher eingegangen.

Das "Flexibilisierungsmodell"[6] sucht eine optimale Lage und Verteilung der Arbeitszeiten. Dazu werden für eine Filiale i die Anteile der Arbeitszeit an der Wochenarbeitszeit, die in Tageshälfte t fallen, mit z_{it} als Variablen eingeführt. Als Information ist die zugehörige Häufigkeit von Produkt j in jeder Tageshälfte für jede Filiale bekannt ($A = (a_{ijt})_{1\leq i\leq m\,;1\leq j\leq n\,;1\leq t\leq T}$ für insgesamt T Tageshälften).

Nach Umformulierungen in den Nebenbedingungen erhält man ein Modell, das die bezüglich Flexiblisierungs-LP optimale Mitarbeiterzahl in jeder Filiale zu jeder Tageshälfte angibt. Über einen Vergleich mit den Istwerten

[4] [BHHH95] S.5

[5] ebenda S. 6f.

[6] ebenda S. 7ff.

ermittelt man den Flexibilisierungsanteil für jede Filiale. Wie beim Modifizierten Modell stehen auch hier Benchmarking-Aussagen und Leistungsvergleiche nicht im Mittelpunkt.

2.3 Einsatz der Data Envelopment Analysis

In den Abschnitten E, F und G[7] benutzen BHHH Ideen, ähnlich wie sie der DEA zu Grunde liegen, um zu weiteren Effizienzaussagen zu kommen. Bevor darauf eingegangen wird, wird nun die DEA-Methode grundsätzlich vorgestellt.

2.3.1 Grundgedanke

Der Begriff der Data Envelopment Analysis geht auf den grundlegenden Aufsatz[8] von Charnes, Cooper und Rhodes aus dem Jahre 1978 zurück und stellt einen mittlerweile etablierten Forschungszweig des Operations Research dar. Die DEA dient dazu, die effizienten DMUs[9] herauszufinden und gleichzeitig für die ineffizienten ein Maß der Ineffizienz anzugeben. Diese Betrachtung erfolgt mit einer stark technologischen Sicht, das heißt, die Produktionstheorie und somit der funktionale Zusammenhang zwischen den Inputfaktoren und den damit erzielten Outputs spielt eine wesentliche Rolle. Annahmen über die Technologie, wie zum Beispiel das Voraussetzen eines bestimmten Verhaltens der Skalenerträge, werden im DEA-Verfahren mit den tatsächlich beobachteten In- und Outputs zusammengebracht und die DMUs mit relativer Effizienz identifiziert. Bemerkenswert ist, dass die Produktionsfunktion selbst dabei gar nicht formuliert werden muss. DEA lässt sich als "nicht-parametrisches Verfahren zur empirischen Schätzung einer sogenannten best-practice Produktionsfunktion"[10] klassifizieren.

2.3.2 Vorgehensweisen der DEA

Kehren wir zu der Definition der Produktivität (hier bezeichnet mit ϑ) als Output geteilt durch Input zurück und formulieren allgemein für n Outputs

[7] ebenda S. 10 ff.

[8] [CCR78]

[9] Der in der englischsprachigen DEA-Literatur geprägte Begriff "decision making unit" (DMU) zur Bezeichnung einer Filiale hat sich auch in der deutschsprachigen Literatur eingebürgert. In dieser Arbeit werden deshalb im DEA-Kontext Filialen auch mit DMU bzw. "Entscheidungseinheiten" bezeichnet.

[10] [DA99] S. 412, vgl. auch [Sch96] S. 168

(y_r) und k Inputs (z_i):[11]

$$\vartheta = \frac{\sum_{r=1}^{n} x_r y_r}{\sum_{i=1}^{k} v_i z_i} = \frac{x_1 y_1 + x_2 y_2 + \cdots + x_n y_n}{v_1 z_1 + v_2 z_2 + \cdots + v_k z_k} \tag{1}$$

DEA gesteht jeder DMU zu, dass sie den Gewichtungsvektor (x_r: Gewichtung des r-ten Outputs; v_i: Gewichtung des i-ten Inputs) so festlegt, dass sie selbst ein maximales ϑ erzielt und damit in einem möglichst günstigen Licht erscheint. Dabei wird in Kauf genommen, dass die betrachtete DMU "unrealistische" Werte wählt. Dieser Nachteil wird durch die sehr starke Aussage im Falle der Nichteffizienz der DMU ausgeglichen. Stellt sich nämlich heraus, dass die DMU sogar bei einem für sie günstigen Gewichtungsvektor nicht effizient ist, kann der Vorwurf, nur die fremdbestimmten Gewichte wären die Ursache, nicht mehr gehalten werden. Die Stärke von DEA kommt demnach am besten bei äußerst umstrittenen oder nicht monetär bewertbaren Gewichten zum tragen. Die obige zu maximierende Größe ϑ würde zu einer nichtlinearen Optimierungsaufgabe führen, deshalb werden die DEA-Modelle regelmäßig in äquivalente primale LP überführt bzw. duale Varianten hinzugezogen.[12]

Schritte in der DEA

1. **Annahmen über die Technologiemenge:** Bevor Aussagen über die Effizienz gemacht werden können, ist es immer notwendig, eine Technologiemenge zu spezifizieren, die dafür als Bezugspunkt dienen soll. Die Technologiemenge drückt aus, „welche Transformations*möglichkeiten* der Unit überhaupt zur Verfügung stehen."[13] Die Menge

   $$\mathcal{T} := \{(z,y)^T \mid \text{Outputvektor } y \text{ kann von Inputvektor } z \text{ produziert werden}\}$$

   heisst Technologiemenge. Hier sind die Annahmen über die Skalenerträge zu machen. Die in der Praxis bedeutendsamsten Technologie $\mathcal{T}^{VRS}$ mit variablen Skalenerträgen[14], wie sie auch bei BHHH angenommen wird, erhält man über die Aussage zur Konvexität der Technologiemenge. Dadurch wird bestimmt, dass positive konvexe Linearkombinationen zweier realisierter Input-Output-Transformationen in der Technologiemenge enthalten sind. $\mathcal{T}$ ist konvex, wenn gilt:

   $$\begin{pmatrix} z_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} z_2 \\ y_2 \end{pmatrix} \in \mathcal{T} \;\Rightarrow\; \left[\lambda \begin{pmatrix} z_1 \\ y_1 \end{pmatrix} + (1-\lambda) \begin{pmatrix} z_2 \\ y_2 \end{pmatrix}\right] \in \mathcal{T} \text{ für } \lambda \in \;[0;1].$$

[11] vgl. [CST00] S. 15

[12] s. [CCR78] S. 431 f., [BCC84] S. 1083

[13] [Sch00] S. 37

[14] engl. variable returns to scale, abgekürzt VRS

Die Technologiemenge lautet für m DMUs:

$$\mathcal{T}^{VRS} := \left\{ \begin{pmatrix} z \\ y \end{pmatrix} \in I\!R^{k+n} \;\middle|\; \begin{pmatrix} z \\ y \end{pmatrix} = \sum_{i=1}^{m} \lambda_i \begin{pmatrix} z_i \\ y_i \end{pmatrix}; \lambda \in \Lambda^{VRS} \right\}$$

mit $\Lambda^{VRS} := \left\{ \lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)^T \in I\!R_+^m \;\middle|\; \sum_{i=1}^{m} \lambda_i = 1 \right\}$. Damit enthält die Technologie die tatsächlich beobacheten Aktivitäten $(z, y)^T$ sowie die durch Linearkombination entstehende konvexe Hülle.

2. **Hüllmenge (Envelope):** Aus dem vorangegangenen Schritt ergibt sich ein konvexes Polyeder mit maximal m Ecken.

3. **Gestaltung des DEA-LP:** Die Optimierungsaufgabe wird als primales oder duales Problem formuliert. Kernpunkt ist stets die Variabilität des Gewichtungsvektors. Dies wird um den Aspekt der Input- oder Outputorientierung ergänzt. Teil F bei BHHH ist beispielsweise inputorientiert, da bei festgehaltener Produktion (Output) ein minimaler Personalaufwand (Input) gesucht wird.

   Exemplarisch sei hier das von Banker, Charnes und Cooper entwickelte Modell[15] dargestellt.

   Beispiel eines DEA-LP: Inputorientiertes BCC-Modell

   Gegeben seien zwei Matrizen $\mathcal{Z}$ mit k Inputarten und $\mathcal{Y}$ mit n Outputarten, in denen für m DMUs die relevanten Daten enthalten sind:

$$\mathcal{Z} = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ z_{k1} & z_{k2} & \cdots & z_{km} \end{pmatrix} \qquad \mathcal{Y} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nm} \end{pmatrix}.$$

   Das Tupel $(z_i, y_i)^T = \big((z_{1i}, z_{2i}, \dots, z_{ki}), (y_{1i}, y_{2i}, \dots, y_{ni})^T\big)$ der i-ten DMU heisst *Aktivität*.[16] Gegeben sei eine Technologie mit variablen Skalenerträgen $\mathcal{T}^{VRS}$. Für die zu evaluierende DMU_o mit der Aktivität $(z_o, y_o)^T$ aus den Daten $\mathcal{Z}$, $\mathcal{Y}$ entwickelt man das inputorientierte BCC-Modell:[17]

[15] s. [BCC84]

[16] in Anlehnung an das engl. 'activity' (vgl. [CST00] S.42)

[17] vgl. [BCC84] S. 1085

Ausgangspunkt ist zunächst ein nichtlineares Modell. Es gilt für die DMU_o ihre Produktivität durch Veränderung der Gewichtungsvektoren u, v zu maximieren und dabei die für alle DMUs geltenden Nebenbedingung (3) einzuhalten.

$$
\begin{aligned}
(BCC_o^{NLP}) \quad \max \quad & \frac{x_1 y_{1o} + x_2 y_{2o} + \cdots + x_n y_{no} - x_0}{v_1 z_{1o} + v_2 z_{2o} + \cdots + v_k z_{ko}} && (2)\\
\text{u.d.N.} \quad & \frac{\sum_{r=1}^{n} x_r y_{rj} - x_0}{\sum_{i=1}^{k} v_i z_{ij}} \leq 1 \qquad (j = 1, \cdots, m) && (3)\\
& v_1, v_2, \cdots, v_k \geq 0 \\
& x_1, x_2, \cdots, x_n \geq 0 \\
& x_0 \in I\!R && (4)
\end{aligned}
$$

Durch Umformung erhält man das äquivalente[18] LP für die betrachtete DMU_o. Durch die Nebenbedingung (6) wird deren Input auf 1 gesetzt, die Variablen x_r, v_i und x_0 müssen so gewählt werden, dass die tatsächlich beobachtete Produktivität aller DMUs (in (7)) eingehalten wird:

$$
\begin{aligned}
(BCC_o^{LP}) \quad \max \quad & x_1 y_{1o} + x_2 y_{2o} + \cdots + x_n y_{no} - x_0 && (5)\\
\text{u.d.N.} \quad & v_1 z_{1o} + v_2 z_{2o} + \cdots + v_k z_{ko} = 1 && (6)\\
& \sum_{r=1}^{n} x_r y_{rj} \leq \sum_{i=1}^{k} v_i z_{ij} + x_0 \quad (j = 1, \cdots, m) && (7)\\
& v_1, v_2, \cdots, v_k \geq 0 \\
& x_1, x_2, \cdots, x_n \geq 0 \\
& x_0 \in I\!R && (8)
\end{aligned}
$$

Das zu (BCC_o^{LP}) korrespondierende duale Programm lautet:[19]

$$
\begin{aligned}
(BCC_o^{DLP}) \qquad min \quad & \pi && (9)\\
\text{u.d.N.} \quad & \pi z_o - \mathcal{Z}\lambda \geq 0 && (10)\\
& \mathcal{Y}\lambda \geq y_o && (11)\\
& e\lambda = 1 && (12)\\
& \lambda \geq 0. && (13)
\end{aligned}
$$

[18] s. [CST00] S. 88

[19] vgl. [BCC84] S. 1084

$e = (1, \ldots, 1)$ ist ein Zeilenvektor mit m Einsen. Die Konvexität der Technologiemenge wird durch Nebenbedingungen ((12) und (13)) sichergestellt. Ziel ist es, eine Inputminimierung ((9) i. V. m. (10)) für DMU_o zu erreichen und gleichzeitig die Restriktionen zum Outputniveau (11) einzuhalten. Aus dem Dualen Programm erklärt sich nun auch das vorzeichenunbeschränkte x_o (vgl. oben Nebenbedingung ((4) und (8)), das durch Dualisieren der Konvexitätsbedingung (12) entsteht.

Ein ausführlicher Überblick über DEA-Modellierungsansätze findet sich zum Beispiel bei [CST00].

4. **Lösen der Optimierungsaufgabe/Bildung Peer Group:** Für jede einzelne DMU wird das DEA-Verfahren angewendet und die Gruppe der effizienten DMUs ermittelt. Dabei darf jeweils die betrachtete DMU ihren produktivitätsmaximierenden Gewichtungsvektor finden. Durch Normierung ist in der Regel ein optimaler Wert[20] von 1 das Kriterium für Effizienz. Für jede DMU i ergibt sich nach Lösen des DEA-LP eine Peer Group (auch Reference Set genannt) E_i. Bezogen auf das obenzuletzt vorgestellte BCC-Modell, enthält diese Menge die bezüglich (BCC_i^{DLP}) effizienten DMUs, d. h. die DMUs, die einen optimalen Wert von $\pi^* = 1$ erreichen.

5. **Messung Abstand:** Für dominierte DMUs lässt sich nun der Abstand zum durch die Peer Group bestimmten effizienten Rand ermitteln. Aussagen, wie welche Inputs und Outputs in welchem Umfang zur Ineffizienz der DMU beitragen, können über Abstandsmaße abgeleitet werden.

6. **Weitere Benchmarking-Aussage:** Für dominierte DMUs ist die Benchmark eindeutig nicht erreicht. Für effiziente DMUs ist ein weiterer Aspekt für ein Ranking innerhalb der Peer Group einsetzbar, nämlich wie oft bei jedem einzelnen DEA-Filialergebnis eine bestimmte DMU in der Peer Group erscheint.

2.3.3 Dualität, Dominanz, Umgang mit Exoten

BHHH setzen das DEA-Verfahren nicht in Reinform um, sondern verbinden es jeweils mit weiteren Instrumenten der Linearen Programmierung. Eine direkte Gegenüberstellung z. B. mit dem oben vorgestellten BCC-Modell ist

[20] Effizienzmaße s. auch [Sch00] S. 80ff.

schwierig, es soll aber gezeigt werden, an welchen Stellen die DEA-Philophie eingesetzt wird und wie sich dies in den Abschnitten E bis G zeigt.

Das bereits als Basismodell unter (2.2.1) vorgestellte LP[21] wird nun in **Abschnitt E** im Sinne der DEA abgewandelt und erneut aufgegriffen. Die DMU_i muss nun versuchen, ihr Residuum r_i durch Veränderung der Gewichtungsvektors x verschwinden zu lassen:

$$
\begin{array}{lll}
(LP_E) & min & r_i \\
 & & Ax + r - s = c \\
 & & \mathbb{1}^T r = \mathbb{1}^T s \\
 & & x, r, s \geq 0.
\end{array}
$$

Findet die DMU_i kein x^* mit $r_i = 0$, ist ihr Produktivitätsniveau nicht zu rechtfertigen, und die DMU_i ist ineffizient bezüglich dieses LPs. Mit dem gefundenen x^* ließe sich ein Reference Set für die DMU_i bestimmen, indem man die r_j für alle anderen DMU_j ermittelt und diejenigen identifiziert, deren Residuum verschwindet ($E_i = \{DMU_j| \ \ r_j = 0 \text{ für } Ax^* + r - s = c, j = 1, ..., m\}$). Wie beschrieben, werden diese DEA-Ansätze aber nicht direkt weiterverfolgt. Im Abschnitt E wird auf der Suche nach dem Mitarbeiterüberschuss der DEA-basierte primale Ansatz dualisiert. Dies eröffnet durch geschicktes Ausnutzen der Komplementarität eine Eliminierung der Variablen bis auf eine, so dass letztendlich eine Benchmarking-Aussage über einen Parameter q möglich wird. Da die Mitarbeiterzahl in der betrachteten Filiale gegenüber den aggregierten Daten aller Filialen gemessen wird, ist das Lösen eines LP gar nicht notwendig.

Ein echte DEA-Version erhielte man, wenn man in LP_E die Produktivität der DMU_i (im Sinne von (1)) maximieren würde, hier mit dem Sonderfall eines einzelnen Inputs, der dementsprechend nicht gewichtet würde. Die Messung der Effizienz im Sinne der so beschriebenen Produktivität ist aber im Abschnitt E nicht das Ziel von BHHH. Statt dessen verfolgen sie einen der DEA ähnlichen Ansatz, indem sie nicht die Produktivität betrachten, sondern die Minimierung des Schlupfes r_i beim Arbeitseinsatz. Führt man dies für alle Filialen durch, kommt man zu Aussagen entprechend der DEA-Philosphie: ist die Mitarbeiterzahl bei für die DMU_i günstigster Gewichtung der Outputs zu rechtfertigen?

Im **Abschnitt F** wird, wie bereits in Abschnitt E, eine DMU_i fokussiert und ausgehend von den beobachteten Produktivitäten die Hüllmenge der DEA-Modellierung für die Ermittlung einer fiktiven Filiale ("auf dem effizienten Rand") genützt. Findet sich eine (echte oder) fiktive Filiale, die die

[21][BHHH95] S. 5

betrachtete Filiale trotz DEA-orientierter, für diese Filiale optimale Wahl des Gewichtungsvektors dominiert, ist diese Filiale nicht effizient.

Um mit Exoten gerechter umzugehen, aber auch um "Mißbrauch" des DEA-Verfahrens durch schwache Filialen zu erschweren, wird im **Abschnitt G** ein gewisser Spielraum für den Gewichtungsvektor eingeführt. Dies wird durch Nebenbedingungen mit Ober- und Untergrenzen für die Gewichtungsvektoren modelliert. Dieses Verfahren wird in der DEA-Literatur mit Assurance-Region-Methode bezeichnet,[22] die 1986 von Thompson et al. [23] entwickelt wurde. Es erlaubt a-priori-Wissen über die Gewichtungsvektoren in die Modellierung mit einzubeziehen.

2.4 Weitere vergleichende Aspekte

Von den einzelnen Methoden lässt sich nicht sagen, welche "besser" oder "schlechter" ist. Vor- und Nachteile ergeben sich unmittelbar aus den jeweils verfolgten, unterschiedlichen Fragestellungen. Für eine eingehende inhaltliche Analyse bringt die Verbindung mehrerer Einzelergebnisse aus beiden Methodenbereichen sicher eine fundiertere Erkenntnis, die somit besser kommuniziert werden kann und sich auch operativ leichter umsetzen lässt. In den folgenden Unterabschnitten werden noch weitere vergleichende Aspekte aufgezeigt.

2.4.1 Benchmarking-Aspekte

Ursprünglich ist Benchmarking der Vergleich von Dienstleistungen über *mehrere* Unternehmen hinweg, bei dem "Unterschiede zu anderen Unternehmen offengelegt" und die "Ursachen für die Unterschiede aufgezeigt werden".[24] Da hier nur innerhalb *eines* Unternehmens Filialen verglichen werden, liegt die einfachste Form des Benchmarkings vor, das interne Benchmarking. Diesen Filialen wird durch die Modellannahmen auch eine äußerst hohe Homogenität hinsichtlich der Struktur der Input- und Outputgrößen unterstellt. Der Benchmarking-Ansatz, der sich bei allen eingesetzten Modellen ergibt, fokussiert nicht die Ursachenanalyse, sondern sucht primär nach Leistungsunterschieden.

[22] zum Beispiel in [CST00] S. 152ff. beschrieben

[23] [TSTS86]

[24] [Hor93] S. 53

2.4.2 Klassifizierung des DEA-Ansatzes

Die Modellierung bei BHHH setzt nichtmonetäre Input- und Outputgrößen ein und ist der ersten Phase der Entwicklungsgeschichte der DEA zuzuordnen. Diese Phase konzentriert sich auf eine Effizienzbetrachtung mit Fokus auf die Technik / Technologie und die Produktion; Marktpreise und Kosten spielen noch keine Rolle.[25] Das Einsetzen der Assurance-Region-Methode ist dabei nur eine Verfeinerung der ersten Phase.

2.4.3 Dominanz: Basis-LP vs. DEA-Modellierung

Durch das Optimieren mit dem Basis-LP wird ein Bewertungsvektor x gefunden, der zugleich auch "sehr produktive" Filialen (positive Komponente(n) in s) und "weniger produktive"(positive Komponente(n) in r) feststellt. Eine Aussage zu Dominanz und Effizienz läßt sich daraus zwar auch formulieren, diese ist jedoch grundlegend verschieden von einer DEA-gestützten Betrachtung: Effizient bezüglich Basis-LP ist die Filiale mit der größten Produktivität $p_i = \frac{A_{i.}x}{c_i}$; da p_i eine skalare Größe ist, dominiert das beste Ergebnis alle Filialen mit $p_j < p_i$. Eine Effizienzbetrachtung mittels DEA würde über die freie Wählbarkeit des Gewichtungsvektors auf bis zu n Dimensionen zurückgreifen. Der einzelnen DMU bietet dies mehr Spielräume, sich dominant zu erweisen, d.h. auf den effizienten Rand der Technologiemenge zu gelangen.

2.4.4 Effizienzranking

Unter Effizienzranking ist die Bildung einer Reihenfolge der DMUs nach dem Sortierkriterium Ineffizienz zu verstehen, ausgedrückt durch den jeweiligen optimalen Zielfunktionswert, das Effizienzmaß.[26] Bei der Interpretation der Ergebnisse ist nochmals in Erinnerung zu rufen, dass die DEA auf die ineffizienten Einheiten fokussiert ist, sie sucht also insbesondere keine *best-practice*, wie dies im Benchmarking-Verfahren gemacht wird. Die effizienten DMU können untereinander grundsätzlich nicht in ein Ranking gebracht werden, es ist allerdings möglich, weitere Kriterien hinzuzuziehen, indem man beispielsweise bei den effizienten DMUs die Häufigkeit im Auftreten der Reference Sets betrachtet.

Für das Ranking der ineffizienten DMUs sei betont, dass die vielleicht naheliegende Aussage „DMU x ist ineffizient, aber noch 'effizienter' als DMU y" nicht richtig ist. Der erreichte Zielfunktionswert sagt nur etwas darüber

[25] Zur Klassifizierung der Entwicklungsphasen der DEA s. [Sen03] S. 11f.

[26] vgl. [Sch00] S. 132

aus, wie das "Verbesserungspotential eingeschätzt"[27] wird.

2.4.5 Best-practice

In der Anwendung der beiden Methoden zeigt sich ein Unterschied in der Berücksichtung der ineffizienten Filialen für eine Best-practice-Betrachtung. Bei der klassischen Linearen Optimierung wie im Abschnitt A gehen über die Mittelwertbildung auch die ineffizienten Filialen ein. Die gefundenen Gewichtungsvektoren sind von allen Filialen gleichermaßen mitbeeinflusst, die best-practice-Filiale ergibt sich wie oben beschrieben aus einer skalaren Größe.

DEA-basierte best-practice DMUs sind über die Logik der Hüllmenge abgebildet und umhüllen die ineffizienten. Die Anzahl der ineffizienten DMUs und deren exakte Lage innerhalb des Envelopes lässt die best-practice DMUs unbeeinflusst.

3 Schlussbemerkung

Die DEA bietet sich als gerade für die Beurteilung nicht monetär bewerteter In- und Outputs, die einer unbekannten Transformationsfunktion unterliegen, als leistungsfähiges Modell an. Dennoch ist die Aussagekraft eingeschränkt, da letztlich alles um die Rechtfertigbarkeit eines Gewichtungsvektors kreist. Die Konzentration von DEA auf ineffiziente Filialen scheint auf den ersten Blick als ausreichend, weitergehende Optimierungsansätze erhalten aber erst durch die Klassische Lineare Programmierung das volle Augenmerk. Insofern existieren diese beiden Methoden auch zurecht innerhalb eines Aufsatzes und ergänzen sich gegenseitig in ihren Fragestellungen.

[27] [Sch00] S. 133

Literatur

[ABHH96] Brita Albers, Achim Bachem, Rainer Heesen, and Winfried Hochstättler. Benchmarking und Leistungsvergleich in Filialnetzen großer Unternehmen. *Zeitschrift für Betriebswirtschaft*, 66(197):1313–1328, 1996.

[BCC84] R. D. Banker, A. Charnes, and William W. Cooper. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30:1078–1092, 1984.

[BHHH95] Achim Bachem, Brita Hecht, Rainer Heesen, and Winfried Hochstättler. Benchmarking und Leistungsvergleich in Filialnetzen großer Unternehmen. http://www.zaik.uni-koeln.de/ paper/preprints.html?show=zpr95-197, 1995.

[CCR78] A. Charnes, W.W. Cooper, and E. Rhodes. Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2(6):429–444, 1978.

[CST00] W. W. Cooper, L. M. Seiford, and K. Tone. *Data Envelopment Analysis*. Kluwer Academic Publ., Boston, 2000.

[DA99] Harald Dyckhoff and Katrin Allen. Theoretische Begründung einer Effizienzanalyse mittels Data Envelopment Analysis (DEA). *Zeitschrift für betriebswirtschaftiche Forschung*, 51(5):411–436, 1999.

[Hor93] Péter Horváth, editor. *Großes Controllinglexikon*. Vahlen, München, 1993.

[Sch96] Michael Schefczyk. Data Envelopment Analysis. *Die Betriebswirtschaft*, 56:167–183, 1996.

[Sch00] Holger Scheel. *Effizienzmaße der Data-Envelopment-Analysis*. Dissertation, Universität Dortmund, Dortmund, Aug. 2000.

[Sen03] Jati K. Sengupta. *New Efficiency Theory*. Springer, Berlin, Heidelberg, 2003.

[TSTS86] R.G. Thompson, F.D. Singleton, R.M. Thrall, and B.A. Smith. Comparative Site Evaluations for Locating a High-Energy Physics Lab in Texas. *Interfaces*, 16:35–49, 1986.

Erklärung

Ich habe die vorliegende Arbeit selbständig verfasst. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlichen oder sinngemäßen Zitate und Entlehnungen deutlich als solche gekennzeichnet.

Pirmasens, im Oktober 2005