

Benchmarking-Ansätze zum Vergleich der Effizienz von Energieunternehmen

**Oliver Franz
Marcus Stronzik**

Nr. 262

Februar 2005

**WIK Wissenschaftliches Institut für
Infrastruktur und Kommunikationsdienste GmbH**

Rhöndorfer Str. 68, 53604 Bad Honnef

Postfach 20 00, 53588 Bad Honnef

Tel 02224-9225-0

Fax 02224-9225-63

Internet: <http://www.wik.org>

eMail info@wik.org

[Impressum](#)

In den vom WIK herausgegebenen Diskussionsbeiträgen erscheinen in loser Folge Aufsätze und Vorträge von Mitarbeitern des Instituts sowie ausgewählte Zwischen- und Abschlussberichte von durchgeführten Forschungsprojekten. Mit der Herausgabe dieser Reihe bezweckt das WIK, über seine Tätigkeit zu informieren, Diskussionsanstöße zu geben, aber auch Anregungen von außen zu empfangen. Kritik und Kommentare sind deshalb jederzeit willkommen. Die in den verschiedenen Beiträgen zum Ausdruck kommenden Ansichten geben ausschließlich die Meinung der jeweiligen Autoren wieder. WIK behält sich alle Rechte vor. Ohne ausdrückliche schriftliche Genehmigung des WIK ist es auch nicht gestattet, das Werk oder Teile daraus in irgendeiner Form (Fotokopie, Mikrofilm oder einem anderen Verfahren) zu vervielfältigen oder unter Verwendung elektronischer Systeme zu verarbeiten oder zu verbreiten.

ISSN 1865-8997

Inhaltsverzeichnis

Tabellenverzeichnis	II
Abbildungsverzeichnis	II
Zusammenfassung	III
Summary	IV
1 Einleitung	1
1.1 Benchmarking	1
1.2 Ablauf der Untersuchung	4
2 Effizienz und Produktivität	6
3 Data Envelopment Analysis (DEA)	9
3.1 Skalenerträge und DEA	10
3.1.1 DEA bei konstanten Skalenerträgen (CRS Modell)	12
3.1.2 DEA bei variablen Skalenerträgen (VRS Modell)	14
3.2 Input- oder outputorientierte DEA	16
3.3 Umgang mit der Slack-Problematik	18
3.4 DEA mit Preisinformationen – Allokative Effizienz	19
3.5 Umweltvariablen und DEA	21
3.6 DEA über mehrere Perioden – Malmquist DEA	23
3.7 Anfälligkeit der DEA Technik für Datenfehler	25
3.8 DEA und Energieversorgungsunternehmen	26
4 Stochastische Verfahren	29
4.1 Alternative Spezifikationen der Produktionsfunktion	29
4.2 Traditionelle Ansätze zur Schätzung von Effizienzgrenzen	32
4.2.1 Methode der kleinsten Quadrate (OLS)	32
4.2.2 Corrected Ordinary Least Squares (COLS)	34
4.2.3 Modified Ordinary Least Squares (MOLS)	36
4.3 Stochastic Frontier Analysis (SFA)	37
5 Zusammenspiel von DEA und SFA	43
Annex 1: Preisindex-Formeln (PIN)	47
Annex 2: Technik der linearen Programmierung bei DEA	49
Literaturverzeichnis	51

Tabellenverzeichnis

Tabelle 3.1:	Häufigkeit des Einsatzes wichtiger Input- und Outputvariablen in 20 Benchmarking Studien zu Energieversorgungsunternehmen	28
Tabelle 5.1 :	Die drei Benchmarking-Verfahren im Vergleich	45

Abbildungsverzeichnis

Abbildung 3.1:	Skalenerträge und DEA	11
Abbildung 3.2:	Effizienzmessung bei input-orientierter DEA (CRS)	13
Abbildung 3.3:	Effizienzmessung bei input-orientierter DEA (VRS)	15
Abbildung 3.4:	Effizienzmessung bei output-orientierter DEA (CRS)	17
Abbildung 3.5:	Allokative Effizienz bei input-orientierter DEA (CRS)	20
Abbildung 3.6:	Input-orientierte Malmquist DEA (CRS)	24
Abbildung 4.1:	Methode der kleinsten Quadrate	33
Abbildung 4.2:	Korrektur einer Produktionsschätzung bei COLS und MOLS	35
Abbildung 4.3:	Mögliche Verteilungen der Produktionseffizienz bei stochastischen Verfahren	38
Abbildung 4.4:	SFA Produktionsschätzung	41

Zusammenfassung

Im Stromsektor stehen die Übertragungs- und Verteilungsnetze im Fokus der Regulierungsdebatte, da sie monopolistische Bottlenecks darstellen und darüber hinaus keine wettbewerblichen Alternativen als Effizienzmaßstab vorhanden sind. Das entscheidende regulatorische Grundproblem besteht dabei in der asymmetrischen Informationsverteilung zwischen den Unternehmen und der Regulierungsbehörde. Ein Weg für den Regulierer, objektive Informationen über die Performance der Unternehmen zu generieren, besteht im sog. Benchmarking.

Die zentrale Fragestellung dieser Studie ist, auf welche Methoden sich die Regulierungsbehörde als mögliche Benchmarking Techniken in der Zukunft konzentrieren sollte. Der Schwerpunkt der Analysen wurde dabei auf die Data Envelopment Analysis (DEA) – ein Verfahren der linearen Optimierung – und die Stochastic Frontier Analysis (SFA) – ein stochastisches Verfahren – gelegt. Darüber hinaus wurden auch die Price Index Numbers (PIN) sowie die traditionellen Regressionstechniken Corrected Ordinary Least Squares (COLS) und Modified Ordinary Least Squares (MOLS) betrachtet.

Während diese traditionellen Ansätze zur Schätzung von Effizienzgrenzen durch die SFA dominiert werden, kann hinsichtlich der drei anderen Verfahren keine eindeutige Aussage getroffen werden. Allerdings erscheint auch die PIN im betrachteten Kontext nicht ratsam, da sie mit zu hoch aggregierten Daten arbeitet. Hinsichtlich der Datenverfügbarkeit können Konsistenzprobleme sowohl bei SFA als auch bei DEA zumindest teilweise umgangen werden, wenn auf rein physische Daten zurückgegriffen wird. Bei der Frage der Güte der Daten erscheint zunächst SFA als vorteilhafter, da sie Datenungenauigkeiten explizit berücksichtigt. DEA hingegen reagiert sensibler auf Datenfehler. Ein weiterer Vorteil von SFA ist, dass die Verlässlichkeit bzw. Robustheit der Effizienzergebnisse mittels Konfidenzintervallen einer Überprüfung unterzogen werden kann. Diesem Mangel bei der DEA ist allerdings – z.B. über Bootstrapping – begegnet worden. Ferner besteht bei DEA nicht die Notwendigkeit, ex ante Annahmen über Kosten- bzw. Produktionsstrukturen sowie die Verteilung der Störterme zu treffen.

Grundsätzlich sollten DEA und SFA nicht als zueinander konfliktär betrachtet werden. Im Gegenteil bieten die zwei verfügbaren Methoden eher eine Möglichkeit an, die ermittelten jeweiligen Effizienzergebnisse auf ihre Plausibilität zu untersuchen (Korrelation der Ergebnisse). Allerdings wird davon abgeraten, diese mechanisch in den Regulierungsprozess zu übertragen (z.B. Bestimmung von X-Faktoren bei der Price Cap Regulierung), da dort auch noch andere Unternehmensspezifika eine Rolle spielen. Aufgrund der Monopolsituation bei den Netzbetreibern ist es anzuraten, inputorientierte Ansätze zu verfolgen, die zwar keine Aussagen über die allokativen Effizienz auf der Outputseite zulassen, welche allerdings bei preissetzendem Verhalten auch von geringer Aussagekraft ist.

Summary

Regulation in the electricity sector is mainly targeted at transmission and distribution as they represent monopolistic bottlenecks. Furthermore, no competitive alternatives do exist that might serve as efficiency benchmarks. The crucial issue concerning the regulatory process is the asymmetry of information about production and cost structures between the regulator and grid companies. In this context benchmarking techniques provide for a promising tool in order to generate more objective information about the efficiency performance of the individual network operators. With over 900 companies in Germany, there should be sufficient data available to make these tools viable.

This study tries to answer the question which method out of the pool of the various benchmarking approaches the regulator should concentrate on in the future. The analysis is centred around the Data Envelopment Analysis (DEA) – a linear programming technique – and the Stochastic Frontier Analysis (SFA) – a stochastic approach. Moreover, Price Index Numbers (PIN) as well as the more traditional regression analysis of Corrected Ordinary Least Squares (COLS) and Modified Ordinary Least Squares (MOLS) are tackled.

Out of the class of econometric approaches the two more traditional ones (COLS and MOLS) are systematically dominated by the Stochastic Frontier Analysis. Concerning the remaining three methods, namely PIN, SFA and DEA, no absolute statement can be derived. If it is intended to measure the productivity/efficiency on a company level, PIN can not be recommended because of the level of aggregation of the utilised data. Concerning the question of data availability, DEA as well as SFA can – at least partly – circumvent inconsistency problems by using data input only in physical terms instead of price information. If the quality of data is poor, SFA seems to be superior to DEA since the former explicitly accounts for data noise whilst the latter does not. DEA results seem to be a bit more vulnerable to outliers. Furthermore, the SFA results can be tested regarding their robustness using confidence intervals. The last argument is not an argument against the use of DEA anymore as this problem can be overcome, e.g. via bootstrapping. One clear advantage of DEA over SFA is that there is no necessity of ex ante assumptions concerning cost or production structures.

All in all, both approaches should be seen as complements and not as competitors. The outcome of DEA can be checked against SFA results and vice versa. The correlations between the results might give first hints regarding problems in the data set or systematic errors. Because of the monopolistic situation of grid companies it is recommended to use input oriented approaches. Finally, it should be stressed that the numbers resulting from DEA/SFA should not be used “mechanically” in the regulatory process, e.g. to determine X-factors under a price cap regime. Other company-specific factors also play a role in achieving certain future productivity goals and should therefore not be overlooked.

1 Einleitung

Das entscheidende regulatorische Grundproblem besteht in der asymmetrischen Informationsverteilung zwischen den zu kontrollierenden Unternehmen und der Regulierungsbehörde. Grundsätzlich verfügen die Unternehmen über bessere und genauere Informationen bezüglich ihrer Kosten, Kunden und Märkte als der Regulierer. Sie haben jedoch keine Anreize, diese Informationen zu offenbaren. Vielmehr mag es ihnen vorteilhaft erscheinen, ihre Lage negativer darzustellen, als sie in Wirklichkeit ist. Dies ist vor allem dann zu erwarten, wenn es sich um ein einzelnes Unternehmen handelt, denn in diesem Fall ist die Behörde kaum in der Lage, die „wahre“ Effizienz der Unternehmung zu beurteilen, und wird daher kaum Preis- oder Kostensenkungen durchsetzen können, ohne sich dem Vorwurf auszusetzen, durch die Regulierung werde der Bestand des Unternehmens gefährdet. Die moderne Regulierungstheorie sucht daher nach Mechanismen, die es dem Regulierer ermöglichen, Informationen über die wirtschaftliche Lage der Unternehmen zu gewinnen. Vor allem ist es wichtig, die Produktivitätsentwicklung der Unternehmen in der Vergangenheit einschätzen zu können, um so ihre Preise an allokativ vorteilhafte Niveaus heranzuführen, ohne dabei Anreize zu weiteren Produktivitätsfortschritten zu untergraben. Hierzu bieten sich regulatorische Verträge an, die dem Unternehmen Anreize bieten, seine wahre Kostenstruktur bzw. das in dieser verborgene Kostensenkungspotenzial zu offenbaren. Dazu zählen z.B. die Preisobergrenzenregulierungen, bei denen die Unternehmen zusätzliche Gewinne, die aus Effizienzsteigerungen resultieren, zunächst für einen gewissen Zeitraum behalten dürfen, solange sie ihre Preisobergrenze nicht verletzen. Während der Mechanismus eines solchen Price Cap als eine Form von Anreizregulierung an sich einfach zu implementieren ist, hat die reale Anwendung dieser Instrumente gezeigt, dass es Regulierungsbehörden vor allem schwer fällt, die notwendigen Informationen zu gewinnen.

1.1 Benchmarking

Ein Weg, Informationen zu generieren, auf die das einzelne Unternehmen keinen entscheidenden Einfluss hat und die insofern auch nicht strategieanfällig sind, besteht im sog. **Benchmarking** oder **Yardstick Competition**, einem Vergleichsmarktkonzept, das sich vor allem dann anbietet, wenn die Regulierungsbehörde sich ohnehin mehreren Unternehmen gegenüber sieht oder wenn in anderen Sektoren (oder Ländern) Unternehmen beobachtet werden können, die den zu regulierenden Unternehmen soweit ähneln, dass sie als Maßstäbe eines belastbaren Vergleiches herangezogen werden können. Eine Problematik kann hier beispielsweise darin bestehen, dass Vergleichsunternehmen nicht gefunden werden können, da die Organisation des fraglichen Sektors im Ausland auf anderen Eigentumsformen oder regulatorischen Ansätzen fußt, die einen signifikanten Einfluss auf die Leistungen des Sektors ausüben können. Dann kann die Vergleichbarkeit soweit eingeschränkt sein, dass die

Ergebnisse kaum gesichert genug sein dürften, um hieraus regulatorische Entscheidungen zu entwickeln. Liegen aber ausreichend vergleichbare Unternehmen auf nationaler Basis vor, so stellt Benchmarking einen sinnvollen Ansatz dar, um die Informationsasymmetrie zwischen Regulierungsinstanz und zu kontrollierendem Sektor zu verringern.

Prinzipiell verfolgen die Benchmarking Ansätze die Idee, dass es bei einer genügend großen Anzahl von Unternehmen unwahrscheinlich wird, dass diese sich bewusst oder unbewusst soweit koordinieren, dass z.B. die durchschnittliche Effizienz der betroffenen Industrie hiervon beeinflusst wird. Da die durchschnittliche Performance ihres eigenen Sektors auch den betroffenen Unternehmen nur grob bekannt sein dürfte und alle Unternehmen Anlass haben, ihre eigene Situation möglichst positiv darzustellen, verlagern Benchmarking-Techniken das Problem der asymmetrischen Information vom Regulierer zurück zu den Unternehmen. Denn es ist davon auszugehen, dass in einer der Kartellproblematik nicht unähnlichen Situation, die Unternehmen im Wissen um den regulatorischen Vergleich Anreiz haben, sich relativ effizient zu verhalten, da ihre eigene Performance relativ zu der anderer Unternehmen beurteilt wird. Denn während die Entwicklung eines Aggregats, das sich aus der Leistung aller Unternehmen ergibt, regulatorische Aussagekraft erhält, kann sich die einzelne Unternehmung nicht sicher sein, welche relative Position sie in der Grundgesamtheit einnimmt.

Dabei sind Benchmarkings grundsätzlich für alle Teilbereiche der unternehmerischen Leistungserstellung denkbar: So kann sowohl die produktive Verwendung der Inputs als auch die hierbei entstehenden Kosten betrachtet werden. Ebenso möglich ist eine output- bzw. erlösorientierte Betrachtung bzw. Kombinationen aus diesen beiden, die sich dann der Entwicklung des unternehmerischen Profits widmen.

Generell lassen sich hierbei Ansätze, die sich auf die durchschnittliche Leistung der Unternehmen beziehen, von solchen unterscheiden, die sich an den besten Unternehmen (frontier) orientieren. Aus ökonomischer Sicht ist dabei die Analyse der von den besten Unternehmen gebildeten Produktivitätsgrenze aus zwei Gründen klar zu bevorzugen:

- Orientiert man sich an Durchschnitten, so werden etwa 50 % der Unternehmen schlechtere Werte aufweisen als der Durchschnitt und insofern der Regulierungsbehörde gegenüber in Rechtfertigungsdruck geraten bzw. aufgefordert sein, künftig bessere Werte zu erzielen. Dagegen werden die anderen 50% der Unternehmen nicht einer genaueren Betrachtung unterzogen, obwohl anzunehmen ist, dass auch unter denjenigen Unternehmen, die den Durchschnitt nur knapp geschlagen haben, sich noch solche befinden, die sich unter Umständen deutlich verbessern könnten. Zu bedenken ist auch, dass Unternehmen, die den Durchschnitt schlagen, deshalb noch nicht in jedem Fall unbedenkliche Tarife ausweisen. Die hier verborgenen Produktivitätsdefizite und Diskriminierungspotenziale bieten die Möglichkeit, die Situation sowohl in

allokativer wie in produktiver Hinsicht¹ zu verbessern. Sie lassen sich jedoch nur beheben, wenn man sich zumindest an der „best-practice“ – also den Kostenführern der fraglichen Industrie – orientiert.²

- Zudem haben Unternehmen, die den Durchschnitt erreichen oder sogar schlagen keinen Anreiz, sich weiter zu verbessern, da sie hierdurch keine Vorteile erzielen können. Es ist daher c.p. kaum zu erwarten, dass sich der Durchschnitt im Zeitablauf verbessert; lediglich die Auflagen an die Unternehmen, die den Durchschnitt nicht erreichen, tragen zu seiner Veränderung bei. Da jedoch nicht ausgeschlossen werden kann, dass Unternehmen, die den Durchschnitt übertreffen, sich in künftigen Perioden weniger anstrengen und den Mittelwert so erhöhen, ist selbst dieser Effekt nicht gesichert.

Bezüglich der in der Literatur diskutierten Benchmarking-Verfahren, die den obigen Bedingungen genügen, weil sie dazu geeignet sind, Informationen über die „best-practice“ zu generieren, lassen sich vor allem drei Ansätze unterscheiden:³

- a) **Data Envelopment Analysis (DEA)** – ein Verfahren der linearen Optimierung, bei dem eine einzelne Unternehmung mit einer Vielzahl von anderen Unternehmen verglichen werden kann. Der Ansatz beruht auf der Bestimmung einer optimalen Produktionsmöglichkeitenkurve aus den Daten aller Unternehmen, gegen die dann die relative Position einzelner Unternehmen evaluiert werden kann. Da die optimalen Input- und Outputkombinationen das Ergebnis eines linearen Optimierungsproblems sind, kann die DEA angewendet werden, ohne explizite Produktions- oder Kostenfunktionen formuliert zu haben. Liegen Daten aus mehreren Perioden vor, so erlaubt eine dynamische DEA anhand des Malmquist-Indexes (**Malmquist DEA**) den Vergleich im Zeitablauf, sodass die relative Entwicklung der Produktivität bestimmter Unternehmen aufgezeigt werden kann.
- b) **Corrected Ordinary Least Squares (COLS)** stellen im Gegensatz dazu ein stochastisches⁴ Verfahren dar, bei dem zunächst der Verlauf einer

¹ Produktive Effizienz bedeutet, dass das Unternehmen im Kostenminimum arbeitet, die Produktionsfaktoren mithin optimal eingesetzt werden. Allokative Effizienz bezeichnet die bestmögliche Versorgung der Konsumenten, d.h. es werden keine Monopolrenten abgeschöpft.

² Eine noch stärkere Annäherung an ein wettbewerbliches Ergebnis ist nur zu erzielen, wenn sich die Regulierung völlig von Istwertbetrachtungen trennt und effiziente Einheiten von Grund auf technisch-ökonomisch modelliert werden. Dies geschieht in der Telekommunikation im Rahmen der analytischen Kostenmodelle.

³ Auch wenn die Ansätze hier differenziert betrachtet werden, so kann festgestellt werden, dass die Praxis sie häufig in Kombination einsetzt. So werden die Ergebnisse von DEA Benchmarkings häufig mit stochastischen Frontieranalysen überprüft oder die DEA-Ergebnisse werden gegen bestimmte (Umwelt-) Variablen regressiert, um herauszufinden, welche signifikant für die beobachteten Abweichungen sind.

Regressionsgerade durch die Daten geschätzt wird. Hierzu bedarf es aber nicht nur Annahmen über den funktionalen Zusammenhang der Variablen; vielmehr müssen zudem explizite Verteilungsannahmen bezüglich der statistischen Fehlerterme etc. getroffen werden. Die geschätzte OLS Kurve wird sodann korrigiert (über Veränderungen des Absolutgliedes nach unten/oben verschoben), um der Tatsache Rechnung zu tragen, dass eben nicht die Erwartungswerte der Variablen gesucht sind sondern deren bestmögliche Ausprägungen. Folge dieser Technik ist, dass die individuellen statistischen Residuen der einzelnen Unternehmen als Ineffizienz interpretiert werden, d.h. es bleibt kein Raum für zufällige Schwankungen der Werte.⁵

- c) **Stochastic Frontier Analysis**, (SFA) ein ebenfalls stochastisches Verfahren, dass im Gegensatz zur problematischen Parallelverschiebung der Regressionsgerade bei COLS versucht, die Frontier für eine vorher definierte Kosten- oder Produktionsfunktion explizit zu schätzen. Hierzu wird die statistische Restgröße der Schätzung aufgespalten in einen Term, der weiterhin die allgemeine Unsicherheit der Schätzung wiedergibt und entsprechend sowohl positive als auch negative Werte annehmen kann, und einen weiteren Term, der die auftretenden Ineffizienzen beschreibt. Dieser zweite „Fehlerterm“ kann folglich entweder nur positive Werte (Kostenschätzung) oder negative Werte (Produktionsschätzung) annehmen.

Neben diesen drei Ansätzen gibt es noch die so genannten **Price Index Numbers** (PIN) sowie die **Modified Ordinary Least Squares** (MOLS), die den beiden obigen Kriterien aber nur bedingt genügen. Während die PIN in der Regel nur eine aggregierte Betrachtung zulassen, ist MOLS eine Weiterentwicklung des COLS Ansatzes, der jedoch – streng genommen – keine Ausrichtung an der Frontier beinhaltet.

1.2 Ablauf der Untersuchung

Im nachfolgenden Kapitel wird ein kurzer Abriss über die allgemeinen Probleme der Messung von Effizienz und Produktivität gegeben, der sich eher den technischen Hintergründen widmet. Dabei wird auch auf die PIN eingegangen werden, die sich systematisch aus diesen Betrachtungen zu Indices ergeben. Anschließend werden die anderen oben genannten Verfahren analysiert, um einen Überblick über ihre Funktionsweise und die hiermit verbundenen Vor- und Nachteile zu geben. Dabei

4 Häufig werden die Verfahren auch in parametrische (COLS, MOLS und SFA) und nicht parametrische (DEA) Verfahren unterteilt (siehe z.B. Coelli et al. 1998). Da die meisten Anwendungen der ersten Kategorie jedoch im Bereich der Stochastik liegen, haben wir diese unter dem Oberbegriff der „stochastischen Verfahren“ subsumiert.

5 Die eigentlich stochastische Methode mutiert praktisch zu einem deterministischen Verfahren.

orientiert sich die Analyse vor allem an den Anforderungen eines neu geschaffenen Regulierungsregimes, das nicht die Möglichkeit hat, auf längere Zeitreihen zurückzugreifen, sondern mit relativ wenigen Informationen die relative Effizienz der Unternehmen abschätzen muss.

Es wird dabei evident werden, dass gerade zu Beginn eines Regimes und bei einer ausreichend großen Zahl von Unternehmen sowohl die DEA als auch SFA einen gewissen Vorteil gegenüber den anderen Verfahren haben, weshalb diese beiden Methoden etwas vertiefter diskutiert werden. Neben den Anforderungen des Instruments wird es dabei auch um die Verwertbarkeit der erzeugten Daten gehen. Diesbezüglich lässt sich zeigen, dass die praktischen Erfahrungen anderer europäischer Regulierer mit der Anwendung dieser Instrumente nicht nur positiv sind. Im Gegenteil wird deutlich werden, dass gerade die Versuche einer „mechanischen“ Umsetzung der erzeugten Effizienzwerte in Zielwerte für Produktivitätsfortschritte oder Preissenkungen zu großen Konflikten zwischen den jeweiligen Behörden und den Unternehmen geführt haben. Von einer mechanischen Anwendung des Instrumentariums wird daher abgeraten – vielmehr scheint es geeignet, eine Art **Aufgreifkriterium** zu bilden. So ist beispielsweise vorstellbar, dass Regulierungsbehörden, die DEA einsetzen, sich vor allem mit solchen Unternehmen näher befassen sollten, die neben schlechten Werten aus den DEA Effizienzanalysen auch noch hohe Netzentgelte aufweisen. In diesen Fällen ist davon auszugehen, dass sowohl in produktiver als auch in allokativer Hinsicht Verbesserungspotenzial besteht.

Trotz der Vorteile, die die Verfasser bezüglich DEA und SFA sehen, geht der Bericht auch ausführlich auf die traditionellen Regressionstechniken ein und stellt ihre Anwendungspotenziale dar.

2 Effizienz und Produktivität

Bevor wir die in der Einleitung aufgeführten Benchmarking Ansätze näher diskutieren, soll ein kurzer Abriss über mögliche Indikatoren zur Messung der Effizienz bzw. Produktivität auf Unternehmensebene gegeben werden, um eine fundiertere Grundlage für die später erfolgende Diskussion der mit den verschiedenen Methoden verbundenen Datenerfordernisse zu schaffen. Bezüglich einer gewählten Messgröße lassen sich zwei Bedingungen formulieren: Zum einen sollte sie eine hinreichende Aussagekraft hinsichtlich der unternehmerischen Produktivität besitzen. Zum anderen stellt sich die Frage, inwiefern für die in den Indikator eingehenden Variablen Daten vorhanden sind bzw. sich erheben lassen und welcher Aufwand dabei anfällt.

Im Falle einer Unternehmung, die mit einem Input (x) nur einen Output (y) herstellt, ist die Messung der Produktivität kein Problem. Die Frage, wie effektiv produziert wird, beantwortet der Quotient

$$(2-1) \text{Produktivität} = y / x$$

Sobald es sich aber um einen Produktionszusammenhang handelt, bei dem durch mehrere Inputs mehrere Outputs erzeugt werden, erhöht sich die Komplexität der Analyse der Produktivität. Betrachtet werden muss unter diesen Umständen die sog. **Totale Faktorproduktivität** (TFP) – ein Maß, das beschreibt, wie gut die Unternehmung in der Lage ist, ihre Inputs miteinander zu kombinieren und so ihre Produktion auch über den Effekt einer reinen Erhöhung der Einsatzmengen hinaus zu steigern. Will man unter diesen Umständen die totale Faktorproduktivität der Unternehmung bestimmen, so muss man Inputs und Outputs indizieren, d.h. einen Weg finden, sie untereinander vergleichbar zu machen. Als Gewichte dienen auf der Inputseite häufig die Anteile an den Gesamtkosten, während auf der Outputseite vor allem die Anteile an den Erlösen genutzt werden.

$$(2-2) TFP = \frac{\text{Output Index}}{\text{Input Index}}$$

Damit sind bei mehreren Faktoren und Produkten also grundsätzlich Preisinformationen notwendig, um die totale Faktorproduktivität der Unternehmung zu messen und im Zeitablauf vergleichbar machen zu können.

$$(2-3) \Delta TFP = \frac{\text{Output Index}_t}{\text{Input Index}_t} \bigg/ \frac{\text{Output Index}_{t-1}}{\text{Input Index}_{t-1}}$$

Grundsätzlich gibt es fünf Quellen, aus denen sich die Veränderung der Produktivität von Unternehmen im Zeitablauf speist: ⁶

- Technischer Fortschritt, d.h. aufgrund technologischer Entdeckungen im betroffenen Sektor gelingt es **allen** Unternehmen besser zu produzieren – die Grenze (frontier) bewegt sich.⁷
- Technische Effizienzsteigerungen, d.h. es gelingt **einer** Unternehmung bei konstanten Mengen und Preisen besser, ihre Produktion zu organisieren und sich so der aktuellen Grenze zu nähern.
- Größenvorteile, d.h. eine Änderung des Umfangs der Produktion führt die Unternehmung zu einer Kombination von Inputs und Outputs, die niedrigere Durchschnittskosten verursacht.
- Allokative Effizienz auf der Inputseite, d.h. das Unternehmen passt seinen Inputeinsatz besser an die auf den Faktormärkten herrschenden Preisverhältnisse an.
- Allokative Effizienz auf der Outputseite, d.h. das Unternehmen passt seinen Produktmix besser an die auf seinen Absatzmärkten herrschenden Preise an.

Soll gemessen werden, wie sich die Produktivität im Zeitablauf verändert, sind Indices zu verwenden. Wie bereits erwähnt, ist die Konstruktion von Indices ohne Preisinformationen nicht möglich (Ausnahme Malmquist-Index), weshalb die Literatur hier auch von sog. Price Index Numbers (PIN) spricht. Bereits im vorhergehenden Abschnitt war darauf hingewiesen worden, dass das grundsätzliche Problem der Konstruktion von Indices darin besteht, die einzelnen Indexglieder gegeneinander zu gewichten. Üblicherweise werden in regulatorischen Fragestellungen, die sich mit TFP befassen, Preisinformationen als Gewichte verwendet. Folgende Indices können dabei zur Anwendung kommen, die mit verschiedenen Vor- und Nachteilen verbunden sind:

- Fisher Output-Index,
- Fisher TFP-Index und
- Törnquist Outputindex.

⁶ Damit entsprechen TFP Indices wie in Gleichung (2-3) grundsätzlich den aus der Statistik bekannten Preis- und Mengenindices, je nach dem ob die Mengen oder die Preise zwischen den Perioden variieren.

⁷ Dabei sollte beachtet werden, dass die ökonomische Theorie drei Formen technischen Fortschritts kennt: 1. Neutraler technischer Fortschritt (Hicks neutral), der das Einsatzverhältnis von Kapital und Arbeit nicht verändert und sich ausschließlich in einer Verbesserung der TFP widerspiegelt, 2. Quasi-kapitalvermehrender technischer Fortschritt (Solow neutral) und 3. Quasi-arbeitsvermehrender technischer Fortschritt (Harrod neutral). In den letzten beiden Fällen wirkt der technische Fortschritt einseitig zugunsten eines Produktionsfaktors und trägt so zu einer Veränderung der optimalen Kombination von Kapital und Arbeit bei. Insbesondere im Rahmen der SFA lassen sich die Annahmen so variieren, dass alle Formen des technischen Fortschritts abgebildet werden können.

Alle diese Indices lösen jedoch das grundsätzliche regulatorische Informationsproblem nicht. Während auf der Output-Seite eine ganze Reihe von Daten vorhanden sein dürften (und sei es, dass mit Hilfe der Umsatzsteuerstatistik auf den Absatz geschlossen wird), sind bezüglich der eingesetzten Ressourcen zumeist deutlich weniger Informationen verfügbar. Auch wenn lediglich Indices ermittelt werden sollen, ergibt sich daher die Notwendigkeit, von den regulierten Unternehmen bestimmte Informationen zu erlangen und gegebenenfalls Vorschriften darüber zu erlassen, wie diese die Informationen aufzubereiten haben.

Kann oder will die Behörde diesen Weg nicht beschreiten, so können indirekte TFP Indices eine Lösung darstellen. Ein indirekter Index beruht auf der Überlegung, dass

$$(2-4) \text{ Wert} = \text{Menge} \cdot \text{Preis} \Leftrightarrow \frac{\text{Wert}}{\text{Preis}} = \text{Menge}$$

ist. Sind also Wertgrößen wie Umsätze, Erlöse oder Personalaufwand bekannt, so lässt sich mit einem geeigneten Preisindex auf die jeweiligen Mengen schließen. Allerdings sollte hier beachtet werden, dass Preisindices eben wiederum Aggregationen von Informationen darstellen, weshalb die tatsächliche Situation einer einzelnen Unternehmung nicht mehr vollständig abgebildet wird.

Das eigentliche Problem der Indices aber besteht darin, dass es nicht möglich ist, die gemessene TFP in ihre Bestandteile zu zerlegen. Tiefgehende Informationen über die Gründe, warum bestimmte Firmen ihre Produktivität schneller steigern als andere oder eben Probleme damit haben, sich zu verbessern, sind somit nicht zu erhalten. Aus regulatorischer Sicht kommt einer solchen Zerlegung aber auch und gerade wegen des zwischen allokativer und produktiver Effizienz bestehenden Spannungsverhältnis große Bedeutung zu.⁸ Da die verschiedenen Indices in dieser Ausarbeitung nur eine untergeordnete Rolle spielen, so z.B. bei der Malmquist DEA, wird aus Gründen der Lesbarkeit auf eine formale Darstellung verzichtet. Ein formaler Überblick zu den PIN sowie eine Erörterung der hier nur angedeuteten Vor- und Nachteile ist im Annex 1 zu finden.

⁸ Aufgrund subadditiver Kostenverläufe bei natürlichen Monopolen kann ein Unternehmen nicht zu Preis = Grenzkosten anbieten, da es sonst langfristig einen Verlust erleiden würde. Um allokativen Effizienz zu gewährleisten, müsste der Netzbetreiber subventioniert werden (Kostenerstattung), was wiederum seinen Anreiz, kosteneffizient zu produzieren reduzieren würde.

3 Data Envelopment Analysis (DEA)

Die Data Envelopment Analysis hat zum Ziel, die Effizienz einzelner Unternehmen relativ zu einer Vergleichsgruppe zu messen. Dabei wird allerdings kein Durchschnitt herangezogen, sondern die Leistung der einzelnen Einheiten wird mit einer effizienten Grenze (frontier) verglichen. Bei der DEA handelt es sich um eine nicht-parametrische Methode. In diesem Fall wird ein lineares Optimierungsverfahren herangezogen, um die effiziente Grenze zu berechnen. Im Gegensatz dazu greifen die im Kapitel 4 diskutierten Verfahren auf stochastische Methoden zurück – der Unterschied liegt dann darin, dass die Stochastik auf Hypothesen aufbaut. Dagegen arbeitet die DEA in ihrer Grundform ohne explizite Annahmen über bestimmte Produktions- bzw. Kostenstrukturen.

Hinsichtlich der Anzahl der zur Analyse herangezogenen Unternehmen (Beobachtungen) gilt grundsätzlich für alle im Rahmen dieser Studie beschriebenen Ansätze, dass mit ihr die Güte der Untersuchung zunimmt, da die Aussagekraft größer wird. So macht es z.B. wenig Sinn, einen Effizienzvergleich nur einer einzelnen Unternehmung zu machen. Sie müsste mit sich selbst verglichen werden und wäre daher immer effizient. Diese Beobachtung ist auch bezüglich der im nächsten Kapitel dargestellten stochastischen Verfahren der Regressionsanalyse und der SFA valide. Je größer dort die Anzahl der Beobachtungen ist, desto trennschärfer werden die Untersuchungsergebnisse.

Allerdings hat die Anzahl der beobachteten Unternehmen sowie die Zahl der zu untersuchenden Parameter bei der DEA aufgrund der Eigenschaften der linearen Optimierung eine sensiblere Wirkung auf das individuelle Benchmarking-Ergebnis eines Unternehmens als dies bei den stochastischen Verfahren der Fall ist. Zunächst ist festzustellen, dass mehr Unternehmen in der Optimierung normalerweise dazu führen, dass sich das unternehmensindividuelle Ergebnis tendenziell verschlechtert – dies hat damit zu tun, dass jedes Unternehmen, das der Analyse hinzugefügt wird, eine zusätzliche Restriktion darstellt. Greift diese Restriktion, verändert das fragliche Unternehmen den Verlauf der Frontier. Daraus folgt, dass sich die Effizienzwerte der übrigen Einheiten im relevanten Bereich verschlechtern. Wirkt die Bedingung jedoch nicht restringent, spielt das hinzugefügte Unternehmen also keine Rolle bei der Bestimmung der Grenze, bleiben die Bewertungen aller anderen Firmen konstant. Im Gegensatz hierzu verbessern sich die ermittelten Effizienzwerte im allgemeinen mit der Anzahl der berücksichtigten Inputs und Outputs, da es durch das Hinzufügen von Variablen wahrscheinlicher wird, dass eine bestimmte Unternehmung in gewisser Hinsicht relativ effizient ist. Denn bezüglich jedes Inputs dürfte es einer Unternehmung gelingen, den jeweils maximalen Output zu schaffen – bzw. vice versa. Hieraus kann man folgern, dass immer mehr Beobachtungen (Firmen) vorliegen sollten als das Produkt von Inputs und Outputs. In der Literatur existiert unter dem Stichwort „rule of thumb“ aber auch der Hinweis, die Menge der Beobachtungen sollte die Summe aus Inputs und Outputs in der Optimierung um mindestens das Zwei- bis Dreifache

übertreffen.⁹ Dies soll sicherstellen, dass die Ergebnisse nicht wegen zu weniger Unternehmen oder zu vieler Variablen verzerrt werden.

Eine weitere Daumenregel kann für einen ersten Plausibilitätscheck der DEA-Ergebnisse verwendet werden. Multipliziert man die Zahl der Inputs mit der Zahl der Outputs, so gibt das Ergebnis näherungsweise die Anzahl derjenigen Einheiten an, die sich als effizient herausstellen dürften.¹⁰

Analysiert man zunächst nur die mengenmäßige Verwendung von Inputs und die hierdurch erzeugten Outputmengen, so benötigt die DEA keine Preisinformationen, was vor allem dann, wenn Informationen nur unter Schwierigkeiten beschafft werden können, ein Vorteil ist. Hinzu kommt, dass Modelle existieren, bei denen die erzeugten Ergebnisse nicht von den Größen oder Maßeinheiten abhängen, in denen beispielsweise die Inputmengen gemessen werden. Es ist also unerheblich, ob der Arbeitseinsatz in Std. oder Tagen gemessen wird, jedenfalls solange, wie zwischen beiden Größen ein festes Verhältnis besteht und alle Unternehmen in der Untersuchung ihre Werte in derselben Einheit gemeldet haben bzw. die Daten adäquat aufbereitet wurden.

Ein weiterer Vorteil der DEA besteht darin, dass ihr eine Verallgemeinerung der eindimensionalen Produktivitätsanalyse zugrunde liegt (vgl. Kapitel 2). Will man Produktivitätskennzahlen verallgemeinern, sodass mehrere Inputs und Outputs berücksichtigt werden können, so ergibt sich das bereits aus der Diskussion der PIN bekannte Problem der Gewichtungen. Während die Gewichte bei Preis- oder Mengenindices exogen bestimmt werden, ermittelt die DEA die Gewichte endogen und zwar so, dass der Quotient aus gewichteten Out- und Inputs maximiert wird.¹¹

3.1 Skalenerträge und DEA

In ihrer Grundform vergleicht die DEA alle Unternehmen einer Datengruppe miteinander – dies beinhaltet die implizite Annahme, dass alle Unternehmen dieselbe Größe haben (siehe unten Abschnitt 3.1.1). Wie Abbildung 3.1 a) für den Fall eines Inputs und eines Outputs zu entnehmen ist, liegt die effiziente Grenze dann auf einem Fahrstrahl aus dem Ursprung des Koordinatensystems, dessen Steigung durch den Produktionspunkt der Unternehmung V bestimmt wird. Diese Firma hat die höchste Produktivität (arbeitet mit den niedrigsten Durchschnittskosten)¹², während sich die anderen Unternehmen W bis Z durch niedrigere (höhere) Kennzahlen auszeichnen. Die Analyse ignoriert jedoch

⁹ In der Literatur firmiert dieses Problem auch unter dem Begriff „course of dimensionality“, das besagt, dass die erforderlichen Daten exponentiell mit den Modellparametern ansteigen (vgl. z.B. Cooper et al. (2000:252)).

¹⁰ Vgl. Coelli et al. (1998: 133ff.).

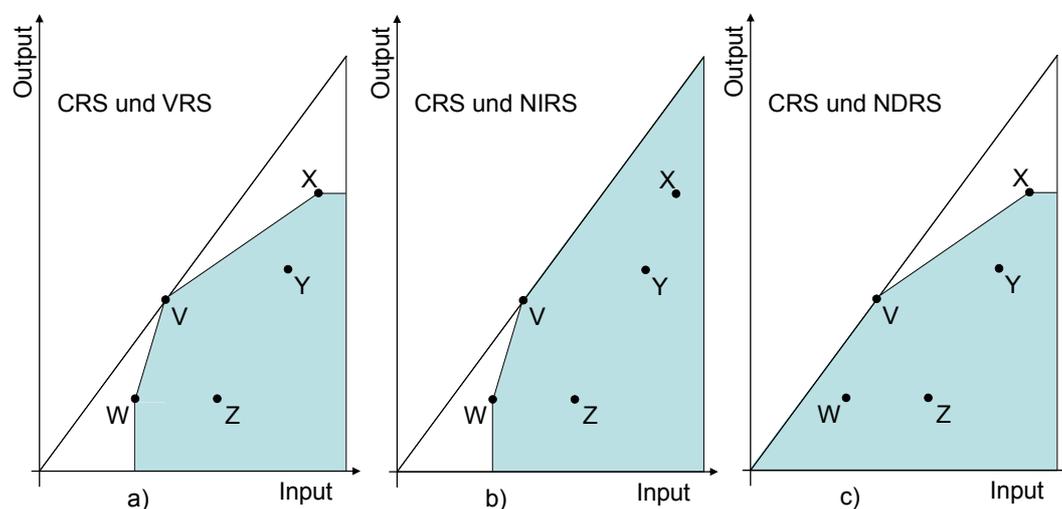
¹¹ Vgl. hierzu auch Annex 2, der einen Überblick über die grundlegende Technik der linearen Optimierung bei der DEA enthält.

¹² Wegen der Dualität von Produktions- und Kostenfunktionen ist es bei ausreichend definierten Produktionsannahmen möglich, direkt aus der Produktionsfunktion auf den Verlauf der Kostenfunktion zu schließen. So bedingen konstante Skalenerträge über den gesamten Bereich der relevanten Nachfrage konstante Grenz- und Durchschnittskosten, wogegen fallende Skalenerträge zu steigenden Grenz- und Durchschnittskosten führen würden.

die Möglichkeit, dass einzelne Unternehmen Nachteile in der Produktion haben, weil sie nicht die optimale Betriebsgröße erreichen oder diese bereits überschritten haben. Diese Größen- oder Skaleneffekte können das Produktionsergebnis wesentlich beeinflussen, was im DEA Grundmodell, dem sog. CRS Modell des nächsten Abschnitts, jedoch nicht berücksichtigt wird.

Weiterentwicklungen des Grundmodells haben jedoch dazu geführt, dass die DEA auch Technologien mit variablen Skalenerträgen (VRS-Modell) analysieren kann. Wie ebenfalls Abb. 3.1 a) zu entnehmen ist, wird die effiziente Grenze in einem solchen Fall abschnittsweise linear und berücksichtigt damit sowohl die relative Effizienz des kleineren Unternehmens W wie auch diejenige des größeren Unternehmens X. W arbeitet in einem Bereich steigender Skalenerträge und müsste folglich seine Produktion ausdehnen, um V zu erreichen, wogegen X quasi schon zu groß ist (fallende Skalenerträge) und seine Produktion einschränken sollte. Wie die Abb. 3.1 b) und 3.1 c) zeigen, ist es zudem möglich, die Skalenannahmen alternativ so zu spezifizieren, dass lediglich fallende (non increasing returns NIR) oder steigende Größenvorteile (non decreasing returns NDR) unterstellt werden.

Abbildung 3.1: Skalenerträge und DEA



Aus ökonomischer Sicht kommt es letztlich auf die Homogenität der zu untersuchenden Gruppe an: Bestehen hinsichtlich der Frage, ob und inwieweit Größenvorteile ausgeschöpft werden, Zweifel, so können diese Effekte quantifiziert werden, wenn ein Vergleich der Ergebnisse unter verschiedenen Annahmen durchgeführt wird. Weichen die ermittelten Werte nicht wesentlich von einander ab, so spielen „scale economies“

keine bedeutende Rolle. Beachtet werden sollte auch, dass die Entscheidung, mit welcher Skalenannahme die Untersuchungen durchgeführt wird, vor allem davon abhängig gemacht werden sollte, inwieweit die Unternehmen über ihre Größe selbst entscheiden können. Sind Fusionen Teil des Entscheidungsspielraumes des Managements und operieren die Unternehmen trotzdem unterhalb der optimalen Größe, so sollten die resultierenden Ineffizienzen den betroffenen Firmen angelastet werden.¹³ Liegen jedoch wettbewerbs- oder eigentumsrechtliche Beschränkungen vor, besteht also kein Freiheitsgrad im Hinblick auf die Unternehmensgröße, so dürfte dieser Aspekt bei der Analyse der relativen Effizienz eine Rolle spielen.

3.1.1 DEA bei konstanten Skalenerträgen (CRS Modell)

Zunächst wird hier davon ausgegangen, dass lediglich die mengenmäßige Verwendung von Inputs und die durch sie erzeugten Outputmengen analysiert werden, was grundsätzlich möglich ist, da die DEA keine Preisinformationen benötigt.¹⁴ Das Grundmodell der DEA geht davon aus, dass alle Unternehmen im ihrem Betriebsoptimum, d.h. dem Bereich konstanter Skalenerträge, arbeiten, weshalb sich ein lineares Optimierungsprogramm wie in Gleichung (3-1) ergibt.

$$(3-1) \min \theta \quad \text{s.t.} \quad -y_i + Y\lambda \geq 0; \quad \theta x_i - X\lambda \geq 0; \quad \lambda \geq 0$$

In Gleichung (3-1) werden die Inputmengen optimiert, d.h. der erzeugte Effizienzwert gibt für alle Unternehmen in der Analyse an, mit wie viel Prozent ihrer derzeit genutzten Inputs sie operieren könnten, ohne dass sie ihre Produktionsmenge reduzieren müssten.

¹³ In einem regulatorischen Umfeld mit frei wählbarer Unternehmensgröße hätte die Entscheidung, zu einer Analyse unter CRS also letztlich zur Folge, dass unmittelbar oder mittelbar Druck auf die Unternehmen ausgeübt würde, sich zu größeren Einheiten zusammenzuschließen.

¹⁴ Allerdings kann DEA mit Informationen über Preise neben Informationen zur technischen Effizienz auch allokativer Probleme abbilden. Vgl. Abschnitt 3.3.

Abbildung 3.2: Effizienzmessung bei input-orientierter DEA (CRS)

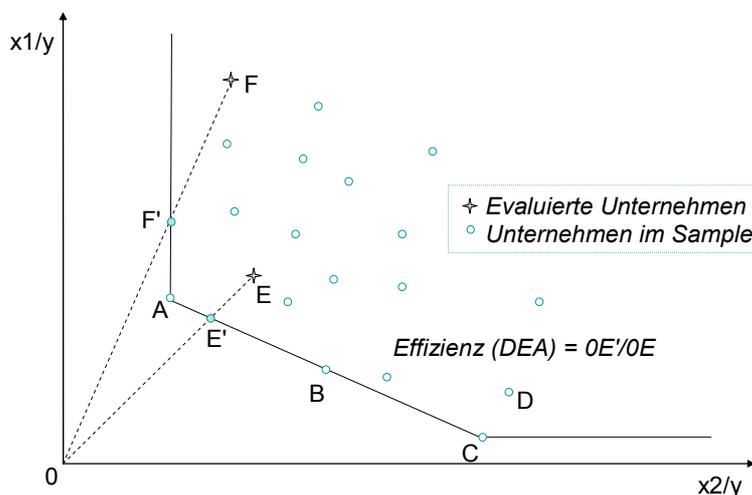


Abb. 3.2 zeigt für den Fall eines Outputs, der mit zwei Inputs produziert wird, eine effiziente Grenze unter der Annahme konstanter Skalenerträge. Die Unternehmen A, B und C bilden diese Grenze – sie liegen auf einer Linie und produzieren **technisch effizient**, wenngleich ihre Faktoreinsatzverhältnisse unterschiedlich sind. Dagegen arbeiten die Unternehmen D, E und F in Abb. 3.2 nicht effizient. So gilt für die Effizienz des Unternehmens E, das sie hinter derjenigen der sog. Peer-Unternehmen (A und B) zurückbleibt.¹⁵ Dies ist ein praktisches Nebenprodukt der DEA, die nicht nur die Peers identifiziert, sondern auch darstellen kann, welche Unternehmen diejenigen Beobachtungen dominieren, die als ineffizient identifiziert werden.¹⁶ Graphisch entspricht der Effizienzwert einem Streckenmaß $\theta_E = OE'/OE < 1$, sodass eine Reduktion der Inputs bei konstantem Output möglich wäre.¹⁷

Betrachtet man nun Unternehmen F, so gilt hier auf den ersten Blick $\theta_F = OF'/OF < 1$. Eine genauere Analyse des auf der Grenze liegenden Punktes F' ergibt jedoch, dass

¹⁵ Korrekt formuliert, wird E mit einer Linearkombination der Unternehmen A und B verglichen, d.h. aus zwei effizienten Unternehmen konstruiert das LP ein virtuelles drittes, das E möglichst ähnlich ist, jedoch effizient arbeitet. Die Anteile mit denen A und B in die Linearkombination eingehen, werden von den meisten DEA Programmen automatisch ermittelt.

¹⁶ Diese Eigenschaft der DEA hat den praktischen Nachteil, dass in der Regulierungsarbeit Unternehmen, die schlecht abschneiden, häufig interessiert sind zu erfahren, mit welchen Konkurrenten sie verglichen wurden.

¹⁷ Vgl. zum Konzept der Distanzfunktionen Coelli et al. (1998: 62-66). Die hier verwendete Größe ist der Kehrwert einer Input-Distanzfunktion – ein Strahlenmaß, d.h. das Verhältnis zweier übereinander liegender Strecken (aus dem Ursprung), beschreibt die „Zielerreichung“ eines bestimmten Produktionspunkts gemessen in Einheiten des Strahls, der die effiziente Grenze erreicht.

das Unternehmen F hier immer noch deutlich mehr an x_1 verbraucht als sein Peer A. Weitere Einsparungen sind also möglich, werden aber von einer DEA in diesem Fall nicht angezeigt, da die zusammengesetzte Grenze parallel zur Achse des betrachteten Faktors verläuft. Die Literatur bezeichnet dieses Phänomen als Input-Slack (im deutschen mit „Schlupf“ zu übersetzen). Grundsätzlich ist es möglich, Inputslack zu identifizieren, denn für eine einzelne Unternehmung i wird nur dann $\theta x_i - X\lambda = 0$ gelten, wenn kein Slack auftritt. Ist die Nebenbedingung jedoch als Ungleichung erfüllt, so ist dies ein Zeichen für das Auftreten von Schlupf. Um das Phänomen genauer analysieren, ist es jedoch notwendig, weitere Berechnungen anzustellen (vgl. Abschnitt 3.3 Umgang mit der Slack-Problematik).

3.1.2 DEA bei variablen Skalenerträgen (VRS Modell)

Die im letzten Abschnitt untersuchten CRS Modelle unterstellen, dass alle Unternehmen die optimale Betriebsgröße erreichen können. Dies mag in der Realität aus mehreren Gründen nicht der Fall sein: Unvollständiger Wettbewerb, Finanzmarktrestriktionen und politische Entscheidungen sind geeignet, Abweichungen vom technischen Betriebsoptimum zu verursachen. Berücksichtigt die Analyse diesen Umstand nicht, so sind die ermittelten Effizienzwerte insofern gestört, als Größennachteile die gemessenen Werte vermindern. Liegen unterschiedliche Unternehmensgrößen vor und ist daher mit dem Auftreten variabler Skalenerträge (variable returns to scale) zu rechnen, so kann dies in einem VRS-Modell wie in Gleichung (3-2) berücksichtigt werden, das technische Effizienz und Skaleneffekte voneinander differenziert.

$$(3-2) \min \theta \text{ s.t. } -y_i + Y\lambda \geq 0; \quad \theta x_i - X\lambda \geq 0; \quad N1'\lambda = 1; \quad \lambda \geq 0$$

Die zusätzliche Nebenbedingung $N1'\lambda = 1$, in der $N1'$ einen Vektor von Einsen bildet, stellt dabei die Konvexität der ermittelten Hülle, wie sie in Abb. 3.3 für den Fall eines Inputs und eines Outputs zu sehen ist, sicher. Diese Bedingung normiert die Summe aller Konstanten λ auf 1 und hat zur Folge, dass ineffiziente Unternehmen nur mit solchen effizienten Beobachtungen verglichen werden, die eine ähnliche Größe haben. Im Gegensatz hierzu vergleicht eine Optimierung unter CRS die Unternehmen unabhängig von ihrer Größe, d.h. wird ein Unternehmen mit größeren Einheiten verglichen, ist die Summe der Konstanten größer im umgekehrten Falle kleiner als 1.

Wie Abb. 3.3 auch zu entnehmen ist, beruht die Unterscheidung von technischer und Skaleneffizienz letztlich auf einem Vergleich der Ergebnisse unter VRS und CRS. Weichen die Ergebnisse nicht voneinander ab, wie bei Unternehmung Q, so liegen konstante Skalenerträge vor. Dagegen fallen die Resultate für Firma R auseinander, so dass Skaleneffekte auftreten müssen. Wiederum ermöglichen es Distanzfunktionen bzw. deren Kehrwerte, die einzelnen Effekte zu beschreiben, wobei die technische Effizienz zweimal, d.h. mit beiden Annahmen bezüglich der Skaleneffekte, ermittelt

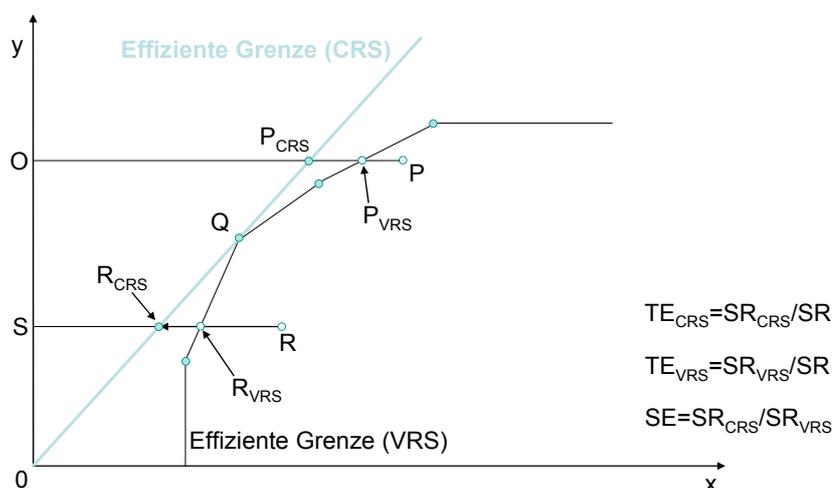
wird. Beiden Maßen ist gemein, dass sie in der Strecke SR normiert sind, weshalb sie in Bezug zueinander gesetzt werden können.

$$\begin{aligned} TE_{CRS} &= SR_{CRS} / SR \\ (3-3) \quad TE_{VRS} &= SR_{VRS} / SR \\ &\Rightarrow SE = SR_{CRS} / SR_{VRS} \end{aligned}$$

Ein Problem dieser Messung von Skaleneffekten besteht darin, steigende nicht von fallenden Skaleneffekten unterschieden zu können, da auch bei Unternehmung P Skaleneffekte angezeigt werden würden, wobei gilt, dass $OP_{CRS} / OP_{VRS} < 1$ ist.

Um hier genauere Resultate zu erzeugen, wird auf die in Abschnitt 3.1 eingeführte Unterscheidung zwischen nicht-fallenden und nicht-steigenden Skalenerträgen zurückgegriffen. Praktisch ist es hierzu nötig, eine dritte DEA durchzuführen, bei der die Restriktion $N'\lambda \leq 1$ (diese Formulierung entspricht nicht-steigenden Skalenerträgen) gesetzt wird, sodass die Unternehmen nur mit gleichgroßen und kleineren Beobachtungen verglichen werden, nicht jedoch mit deutlich größeren Einheiten.

Abbildung 3.3: Effizienzmessung bei input-orientierter DEA (VRS)



Die Art der auftretenden Skaleneffekte lässt sich dann anhand folgender Regel identifizieren:

$$(3-4) \quad \begin{aligned} IRS &= TE_{VRS} \neq TE_{NIRS} \\ DRS &= TE_{VRS} = TE_{NIRS} \end{aligned}$$

Diese Herangehensweise liefert ein Unterscheidungsmerkmal für die Unternehmen, sofern tatsächlich unterschiedliche Skalenerträge auftreten. Eine vorherige Gruppierung und damit verbunden ein beschränkter Vergleichsmaßstab sind nicht notwendig, um derartige Nachteile abbilden zu können.

In Abb. 3.3 arbeitet nur die Unternehmung Q mit konstanten Skalenerträgen. Unter der Annahme U-förmiger Durchschnittskostenfunktionen steht aber zu erwarten, dass mehrere Unternehmen mit unterschiedlichem Output konstante Skalenerträge aufweisen werden. Dann würden die Effizienzgrenzen beider Modelle links und rechts von Q parallel zueinander verlaufen. Denn mehrere effiziente Unternehmen verbreitern die effiziente Zone, wofür die Literatur den Begriff **Most Productive Scale Size** (MPSS) geprägt hat. Die Ausdehnung einer solchen MPSS lässt sich im Rahmen eines abgewandelten LP bestimmen.¹⁸

3.2 Input- oder outputorientierte DEA

Bisher erfolgte die Analyse der DEA als Methode ausschließlich unter der Annahme, dass bezüglich bestimmter Produktionsniveaus untersucht werden sollte, inwiefern der Einsatz der Inputs optimal, d.h. technisch effizient, erfolgte. Es handelte sich insofern um ein Minimierungsproblem. Die mikroökonomische Analyse kennt aber auch Maximierungsprobleme, bei denen für bestimmte Inputs der größtmögliche, technisch effizient produzierte Output gesucht wird. Dementsprechend kann DEA auch mit Blick auf das Produktionsergebnis angewandt werden, wobei wiederum Distanzfunktionen zu betrachten sind. Liegen CRS-Annahmen zugrunde entsprechen sich die Ergebnisse der Effizienzanalyse aus beiden Ansätzen; sie fallen jedoch auseinander, wenn andere Skalenansätze verfolgt werden.¹⁹

Das LP zur Lösung einer outputorientierten DEA unter CRS hat dann die Form von Gleichung (3-5):

$$(3-5) \max \phi \text{ s.t. } -\phi y_i + Y\lambda \geq 0; \quad x_i - X\lambda \geq 0; \quad \lambda \geq 0$$

Wobei gilt, dass $1 \leq \phi < \infty$ ist und der einer analysierten Unternehmung zugewiesene technische Effizienzwert sich aus $1 \leq 1/\phi < 0$ ergibt. Auch die outputorientierte DEA lässt unterschiedliche Annahmen über die Skalenerträge zu, die dann ebenfalls über eine zusätzliche Restriktion in das System eingeführt werden können.

In Abb. 3.4 produzieren die Unternehmen mit jeweils einem Input zwei unterschiedliche Outputs. Dabei definieren die Punkte A, B und C die effiziente Grenze unter der Annahme konstanter Skalenerträge. Dagegen arbeiten die durch D und E

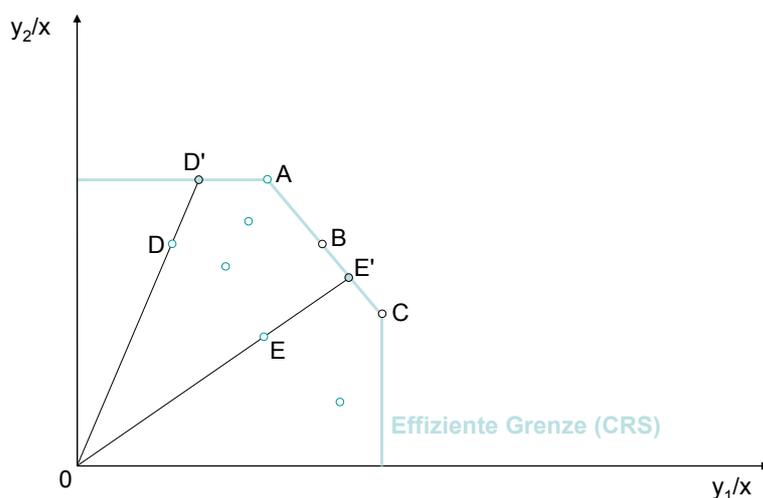
¹⁸ Vgl. Cooper, Seiford, Tone (2000: 127-131).

¹⁹ Vgl. Coelli et al. (1998: 158).

symbolisierten Unternehmen nicht effizient, d.h. sie könnten ihren Output steigern, ohne mehr Ressourcen zu verbrauchen.

Das Unternehmen D bildet insofern eine Ausnahme, als es sich wiederum um einen Fall von Schlupf handelt, denn der Vergleichspunkt D' berücksichtigt nicht, dass im Vergleich mit Unternehmen A der Output an Gut y_1 weiter ausgedehnt werden müsste, um die effiziente Grenze zu erreichen. Wiederum würde die DEA zwar das Vorliegen von Output-Schlupf, aber nicht dessen Umfang anzeigen, da hier analog zur Diskussion der Input-DEA $-\phi y_i + Y\lambda = 0$ gelten muss, um eine Situation ohne Slack zu gewährleisten.

Abbildung 3.4: Effizienzmessung bei output-orientierter DEA (CRS)



Zur Frage, ob Inputs oder Outputs im Zentrum der Analyse stehen sollten, werden unterschiedliche Auffassungen vertreten, die zudem mit dem jeweiligen Untersuchungsgegenstand variieren. Im Hinblick auf Unternehmen der Elektrizitätswirtschaft sind in der Vergangenheit beide Formen der Analyse eingesetzt worden, was auf einen wichtigen Unterschied zwischen der DEA und ökonometrischen Methoden hinweist: Bei DEA ist nicht nötig ex ante eine Beziehung zwischen den Größen zu definieren bzw. die Richtung einer Reaktion festzulegen. Die Frage nach einer Orientierung an Inputs oder Outputs reduziert sich damit auf die Frage über welche Seite des Produktionsprozesses das Management der zu untersuchenden Firmen mehr Kontrolle hat. Im Falle entbundelter Elektrizitätsverteilernetze dürfte kurzfristig mehr Kontrolle über die Inputs gegeben sein, wenngleich darauf zu achten ist, dass diese inputorientierte Analyse der relativen Effizienz nicht negative

Auswirkungen auf die Investitionsanreize hat. Dies gilt vor allem im Hinblick auf die Überwindung von Netzengpässen.²⁰

3.3 Umgang mit der Slack-Problematik

In der Literatur werden zum Umgang mit der Slack-Problematik eine Reihe von Vorschlägen gemacht, die im folgenden kurz vorgestellt werden sollen. Zuvor sei jedoch darauf hingewiesen, dass es sich bei Schlupf letztlich um eine Art Artefakt der gewählten Untersuchungsmethode und ihres Umfangs handelt: Im theoretischen Extremfall einer Untersuchung mit einer unbegrenzt großen Anzahl von Beobachtungen gleicht die effiziente Grenze im Falle zweier Inputs immer mehr einer stetig differenzierbaren Isoquanten, d.h. die Kantenabschnitte werden immer kleiner und gehen im Fall von unendlich vielen Unternehmen in einzelne Punkte über. Unter diesen Umständen ist es dann möglich, für jedes ineffiziente Unternehmen ein Vergleichsobjekt zu finden, ohne dass Input- oder Outputschlupf ein Problem darstellt. Aus diesem Grund ist grundsätzlich für große Untersuchungsgruppen zu plädieren. Die Vorschläge zum Umgang mit dem Schlupfproblem lassen sich grob in zwei Gruppen unterscheiden:

Eine Gruppe von Modellen bezieht Schlupfvariable explizit in die Zielfunktion des zu optimierenden Systems ein, dabei werden die Schlupfvariablen zu den Ungleichungen der Restriktionen addiert. Im sog. **additiven Modell** greifen die Nebenbedingungen dann als Gleichungen:²¹

$$(3-6) \min -(s^- + s^+) \text{ s.t. } -y_i + Y\lambda - s^+ = 0; x_i - X\lambda - s^- = 0; N^1\lambda = 1; s^-, s^+ \geq 0$$

Das additive Modell hat den Vorteil, dass es Input- und Outputschlupf simultan betrachtet. Dabei wird allerdings derjenige Punkt auf der effizienten Grenze gesucht, der möglichst **weit** von der untersuchten Merkmalsausprägung entfernt ist. Dem additiven Modell ist folglich mit anderen Ansätzen, die die Slack-Problematik zu beherrschen suchen, gemein, dass die Schlupfwerte **maximiert** werden, was zur Folge hat, dass nicht der am nächsten gelegene Punkt auf der Frontier gefunden wird sondern der weitest entfernte. Hinzu kommt, dass der vorgestellte Ansatz anfällig für Veränderungen der Maßeinheiten ist – eine Umstellung der Analyse des

²⁰ Zu denken ist hier vor allem an Netzausbauten, die nicht sofort einsetzbar sind. Der Kapitalstock eines Unternehmens, das netto investiert, aber diese Investition nicht innerhalb einer Periode einsetzen kann, müsste dann für die DEA eventuell angepasst werden. In diesem Fall würden DEA und Bestimmung der regulatorisch relevanten Kapitalbasis dann auseinander fallen.

²¹ Vgl. Adler et al. (2002: 252), Coelli et al. (1998: 175) und Cooper et al. (2000: 96f.). Zu der hier vorgestellten Version des Modells, die auf einer VRS-Technologie beruht, existieren zahlreiche Varianten, u.a. das sog. slack-adjusted model und das slack based measure of efficiency. Letzteres ist dazu geeignet, in den gewählten Maßstäben invariante Ergebnisse zu erzeugen. Da hierzu Input- und Outputschlupf jedoch in einen Quotienten gestellt werden, ergibt sich bei den effizienten Unternehmen, deren Slack Null beträgt, ein Problem mit der mathematischen Definition des Ausdrucks.

Arbeitseinsatzes von Tagen auf Stunden würde daher mit einer gewissen Wahrscheinlichkeit zu einer Veränderung der gemessenen Slackwerte führen.²² Dies muss als entscheidender Nachteil der Analyse gelten. Coelli et al. (1998) schlagen im Gegensatz hierzu einen Ansatz vor, bei dem eine mehrstufige DEA durchgeführt wird, wobei der Vorteil darin liegt, dass: "... it attempts to preserve the original input and output mixes as much as possible, and that the method is also invariant to units of measurement."²³ Dabei gehen die Autoren folgendermaßen vor: Sie führen z.B. eine zweistufige input-orientierte DEA (s.o.) durch, um die Unternehmen ohne Slack und mit einem $\theta=1$ von denjenigen unterscheiden zu können, die sich durch Schlupf auszeichnen.²⁴ Hiernach wird für alle Slack-Unternehmen bezüglich aller Inputs erneut eine Optimierung durchgeführt, bei der jedoch nur jeweils einer der Inputs variabel gehalten wird – dies erlaubt zu unterscheiden, bei welchen Inputs das Problem überhaupt auftritt.²⁵ Sodann wird erneut optimiert, wobei jetzt alle Inputs mit Schlupfeigenschaften als optimierbar gelten, wogegen solche, die keinen Slack aufweisen, Teil der Restriktionen sind. Unter bestimmten Umständen müssen diese Abläufe mehrmals wiederholt werden, um den gesamten Schlupf zu eliminieren. Insofern steigt der Rechenaufwand im Vergleich zur zweistufigen DEA deutlich an, andererseits wirkt sich in der praktischen Durchführung positiv aus, dass nur für einen Teil der Unternehmen zusätzliche Durchläufe erforderlich sind.

Es sei zum Abschluss der Diskussion über Slack nochmals darauf hin gewiesen, dass sich die beschriebene Problematik aus der verwendeten Technologie ergibt und zusätzlich mit dem Umfang der Beobachtungen variiert – größeren Gruppen ist daher grundsätzlich der Vorzug vor kleineren zu geben. Zudem lässt sich unserer Ansicht nach argumentieren, dass die Verschwendung von Inputs oder Outputs, wie sie die Schlupfproblematik anzeigt, nichts anderes ist als alloкатive Ineffizienz, denn offenbar werden Faktoren eingesetzt (erzeugt), die ihre Grenzkosten (Wertgrenzprodukte) nicht erwirtschaften.²⁶ Derartigen Abweichungen ist der folgende Abschnitt gewidmet.

3.4 DEA mit Preisinformationen – Allokative Effizienz

Allen zuvor analysierten DEA Ansätzen war gemein, dass sie ohne Preisinformationen durchgeführt wurden. Dies ist einerseits eine Stärke der DEA-Technologie, die vor

²² Zu denken wäre hier insbesondere an ein zweistufiges LP (VRS), dass zunächst optimale Werte für θ ermittelt, um dann unter Verwendung dieser θ als Inputparameter den Schlupfumfang zu optimieren. Vgl. Coelli et al. (1998: 175). Die gleiche Kritik gilt auch für das „slack-based-measure of efficiency (SBM)“ nach Cooper et al. (2000). Zwar kann hier die Problematik der Maßeinheiten beherrscht werden, indem ein Quotient aus Slack- und Variablenwert gebildet wird, jedoch bringt dieses Vorgehen definitorische Probleme bei $x_i=0$ mit sich. Zudem bleibt unklar, warum der Slack mit den Input- oder Outputmengen gewichtet werden sollte.

²³ Vgl. Coelli et al. (1998: 144).

²⁴ Unternehmen, für die kein Slack gefunden wird, obwohl sie sich durch $\theta < 1$ auszeichnen, werden folglich nicht weiter analysiert.

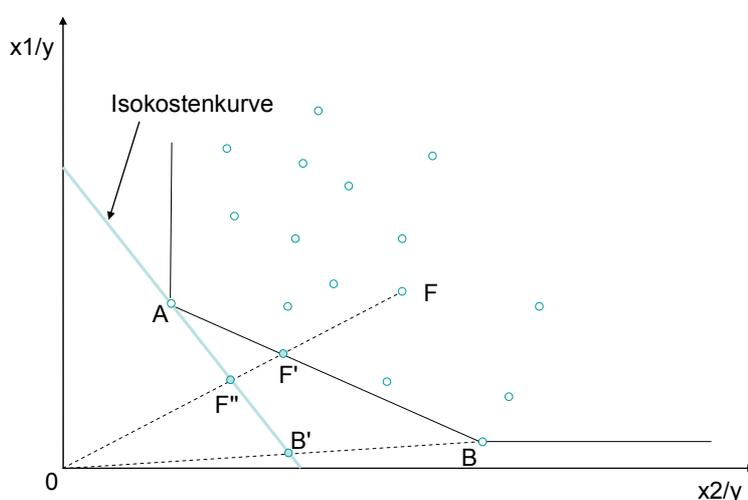
²⁵ Auch diese Lösungsmethode hat Probleme bei Inputs, die nicht verwendet werden, d.h. bei $x_i=0$.

²⁶ Vgl. zu dieser These auch Ferrier und Lovell (1990).

allein dann wirkt, wenn Marktpreise nicht oder nur unvollständig verfügbar sind. Andererseits wurde bereits in Abschnitt 2 Effizienz und Produktivität erläutert, dass es im Rahmen des unternehmerischen Handelns mehr als eine Quelle für Ineffizienzen gibt: Stehen daher Preisinformationen zur Verfügung und kann angenommen werden, dass die Unternehmen einer gemeinsamen Verhaltensannahme genügen (Kostenminimierung oder Umsatz- bzw. Gewinnmaximierung), dann lassen sich allokativen und technische Effizienz gleichzeitig mit Hilfe einer DEA messen. Dabei werden die Inputs und Outputs aller Unternehmen mit den tatsächlichen Güter- oder Faktorpreisen bewertet, d.h. es bedarf keiner einschränkenden Annahmen bezüglich der Wettbewerbssituation auf vor- oder nachgelagerten Märkten. Vielmehr stellt sich neben der Frage, wie viel von einem bestimmten Faktor verwendet wird, auch die Frage, was dieser die betreffende Einheit kostet.²⁷

In Abb. 3.5 wird jedoch davon ausgegangen, dass sich alle Unternehmen denselben Preisen für die Faktoren x_1 und x_2 gegenüber sehen, sodass sich das relative Preisverhältnis p_1/p_2 als Isokostenlinie darstellen lässt.

Abbildung 3.5: Allokativen Effizienz bei input-orientierter DEA (CRS)



Die Isokosten-Kurve schneidet die effiziente Grenze im Punkt A, sodass nur diese Unternehmung sowohl allokativ als auch technisch effizient ist. Dagegen ist der

²⁷ Berücksichtigt man jedoch die Tatsache, dass die zu analysierenden Unternehmen Monopolisten z.B. aufgrund von Bottlenecks in der Netzinfrastruktur sind, so spricht dies unserer Ansicht nach dafür, sich mit Kostenminimierung zu befassen, da Umsatz und Gewinn eben auch zunehmen und besser bewertet werden, wenn ein Unternehmen höhere Preise durchsetzen kann, was im fraglichen Wettbewerbsumfeld jedoch nicht gewollt sein kann.

Produktionspunkt F in beiderlei Hinsicht ineffizient. Es gilt daher wie bisher, dass die technische Effizienz von F durch die Strecke $TE_F = OF'/OF$ wiedergegeben wird. Für die allokativen Effizienz gilt dementsprechend $AE_F = OF''/OF'$. Da implizit angenommen wird, alle Unternehmen hätten zum Ziel, ihre Kosten zu minimieren, bedingen technische und allokativen Ineffizienzen jeweils höhere Kosten als notwendig. Insofern lässt sich die Kosteneffizienz (CE) von F aus der technischen und allokativen Effizienz berechnen.

$$(3-7) CE_F = TE_F \cdot AE_F = OF'/OF \cdot OF''/OF' = OF''/OF$$

Im Gegensatz hierzu entspricht für das Unternehmen B die Kosteneffizienz der allokativen Effizienz, da B in Bezug auf die technische Umsetzung der Produktion als effizient identifiziert wird ($TE_B = OB/OB = 1$). Grundsätzlich lässt sich allokativen Effizienz auch im Hinblick auf Outputs betrachten; bei mehreren Produkten steht dann quasi die Frage im Mittelpunkt, ob der gewählte Preismix umsatz- oder ertragseffizient ist. Grundsätzlich dürfte jedoch die Prämisse gelten, dass die hier zu analysierenden Unternehmen in Teilen über Monopolmacht verfügen und diese auch nutzen – soweit sich der gewählte Preismix eines Unternehmens dann nicht als gewinnmaximal im Sinne einer Abschöpfung von Monopolrenten herausstellen würde, wäre diese „Ineffizienz“ regulatorisch zu verschmerzen.

Ein gewichtiges Argument für die Verwendung einer Kosten DEA – d.h. unter Verwendung von Preisinformationen –, liegt darin, dass bei der reinen Betrachtung von Input- und Outputmengen Outsourcing, Leasing und ähnliche Kostenverlagerungsinstrumente eingesetzt werden können, um die ermittelten Effizienzwerte zu verbessern. Durch das Verlagern von z.B. Service-Leistungen in eine externe Firma, geht der gemessene Arbeits- und Kapitaleinsatz zurück, da weniger eigene Ressourcen benötigt werden, um das Produkt zu erzeugen. Soll dieser Effekt vermieden werden, müssen neben den gesamten Kapitalkosten (z.B. für eine eigene Service-Tochter) auch die alle laufenden Betriebskosten (hierunter fallen die Ausgaben für Outsourcing etc.) berücksichtigt werden, um den Ressourceneinsatz in Relation zur Produktion ökonomisch aussagekräftig zu halten. Dies ist jedoch nur bei Verwendung von Preisinformationen - also Kosten-DEA - möglich.

3.5 Umweltvariablen und DEA

Bei der Durchführung einer DEA wird man häufig auf Variable treffen, auf die das Management der Unternehmen keinen direkten Einfluss hat, die aber dennoch signifikanten Einfluss auf das Produktionsergebnis haben. Hierzu gehören vor allem Umweltvariablen, d.h. im Falle der Verteilung von elektrischer Energie können u.U. die Beschaffenheit des Geländes, die Siedlungsstruktur, die durchschnittliche Kundengröße und ähnliche Aspekte als gegeben und kurzfristig unveränderlich interpretiert werden. Es gibt zahlreiche Möglichkeiten, die DEA so anzupassen, dass derartige

Besonderheiten abgebildet werden können. Ebenso ist es möglich, die Variablen in der Optimierung zu beschränken, sodass sie nur noch bestimmte Höchst- oder Niedrigstwerte annehmen können.²⁸ Jedoch sollte bedacht werden, dass ein derartiges Öffnen der Analyse mittelfristig zu einer Proliferation der Ausnahmen führen wird, denn die Unternehmen haben ein Interesse daran, in einem möglichst günstigen Licht dazustehen. Ausnahmetatbestände anzumelden und durchzusetzen, ist vorteilhaft, da verbesserte Effizienzwerte die Folge sind.²⁹

Grundsätzlich kann die Analyse im Falle einer Berücksichtigung von Umwelteinflüssen dann so erfolgen, dass nicht-diskretionäre Inputs, d.h. z.B. die Geländestruktur, als eine zusätzliche Variable Z in das LP eingeführt werden. Jedoch haben alle Vorschläge, die zu diesem Thema bisher in der Literatur gemacht wurden, Nachteile. So verlangen einige Modellspezifikationen stetige Variablen oder die Wirkrichtung eines Umwelteinflusses muss a priori festgelegt werden.

Ergebnisoffener ist es, die Ergebnisse der DEA wiederum einer ökonometrischen Analyse zuzuführen, um so herauszufinden, ob zwischen schlechtem Abschneiden in der DEA und bestimmten Realisationen von Umweltvariablen ein Zusammenhang besteht. Zusätzlich lässt sich mit Hilfe statistischer Hypothesentests eine Aussage über die Stärke des Zusammenhangs gewinnen. Allerdings ergeben sich gerade bei kleinen Untersuchungsgruppen erneut Probleme: Zunächst dürfte eine solche Regression ohnehin über ihre Freiheitsgerade beschränkt sein. Daneben sollte bedacht werden, dass im Fall einer Analyse mit Umweltvariablen für eine ganze Reihe von Unternehmen ein Effizienzwert von 1 gemessen werden dürfte. Eine Kleinstquadratschätzung wird dann aufgrund der normalverteilten Zufallsvariablen aber auch Effizienzwerte vorhersagen, die größer als 1 sind. Diesem Problem kann auf mehrere Arten begegnet werden:

- Da die Wahrscheinlichkeitsverteilung der DEA-Effizienzwerte im Normalfall unbekannt ist, muss ohne weitere Annahmen (s.u.) auf nicht-parametrische Statistik zurückgegriffen werden. Einen solchen Weg stellt der **Rang-Summen-Test** (nach Wilcoxon, Mann und Whitney) dar, der es erlaubt, die DEA Ergebnisse oder auch nur die effizienten Grenzen zweier Gruppen zu vergleichen.³⁰ Möglicherweise könnte aber auch der Einsatz eines parameterfreien Korrelationskoeffizienten wie z.B. Kendalls Tau, der ursprünglich für ordinalskalierte Merkmale entwickelt wurde, Ergebnisse erzeugen.

²⁸ Ein solches Vorgehen könnte z.B. der Tatsache Rechnung tragen, dass bestimmte Outputniveaus aufgrund von Umweltauflagen nicht erreicht werden können.

²⁹ Dies spricht dafür, neben DEA noch andere Benchmarking Methoden im Regulierungsprozess komplementär einzusetzen, um diese Variablen aus dem Entscheidungsprozess herauszufiltern. Dabei sind neben den im folgenden Kapitel beschriebenen stochastischen Ansätzen – hier insbesondere die SFA – auch analytische Kostenmodelle (Modellnetze) geeignet, wobei letztere zusätzlich die durch die „Netzgeschichte“ bedingten Mehrkosten separieren können.

³⁰ Vgl. Cooper et al.(2000 : 200-205).

- Schätzungen können im sog. **Tobit-Verfahren** durchgeführt werden.³¹ Diese Technik wurde allerdings entwickelt, um ökonometrische Analysen durchführen zu können, obwohl bestimmte Variable den Wert Null annehmen. Die Effizienzergebnisse aus der DEA müssen, um im Tobit-Verfahren eingesetzt werden zu können, zunächst transformiert werden, sodass gilt: $\theta_i^{Tobit} = \theta_i - 1$.
- Zusätzlich lässt sich die **Bootstrap-Methode** einsetzen, die zum Ziel hat, die DEA Werte mit einer Ausgangsverteilung zu hinterlegen, die wiederum eine Aussage über die Wahrscheinlichkeit bestimmter Merkmale erlaubt. Ausgangspunkt der Überlegungen ist dabei die tatsächlich beobachtbare Verteilung der Merkmale. Wobei davon ausgegangen wird, dass diese zumindest repräsentativ für die zugrunde liegende Wahrscheinlichkeitsverteilung und Wahrscheinlichkeitsdichtefunktion ist. So dann gilt es zu entscheiden, welche statistische Eigenschaft analysiert werden soll – im hier vorliegenden Fall sind dies zumeist der Mittelwert und die Standardabweichungen der DEA-Werte. Daher werden in einem Urnenexperiment zahlreiche Ziehungen aus der ursprünglichen Merkmalsverteilung vorgenommen ($n > 1000$). Für jedes dieser Sample werden dann die gefragten statistischen Größen berechnet, die dann wiederum in einem Histogramm abgetragen werden, dessen Mittelwert und Standardabweichung dann als Schätzer der wahren Werte dienen. Mit den so ermittelten Werten lassen sich dann Konfidenzintervalle für die DEA-Werte angeben.³²

3.6 DEA über mehrere Perioden – Malmquist DEA

Zusätzlich zur Frage, wie sich die Unternehmen eines zu regulierenden Sektor in einer bestimmten Periode relativ zueinander stellen, kann auch die Entwicklung der Produktivität über mehrere Perioden wertvolle Informationen für die regulatorische Arbeit liefern. Neben der ökonometrischen Analyse von Panel-Daten, die im folgenden Abschnitt behandelt werden wird, können hierzu auch DEA-artige Techniken verwandt werden – z.B. die sog. Malmquist DEA, die zugleich einen starken Bezug zur Diskussion der PIN in Abschnitt 2 aufweist.³³

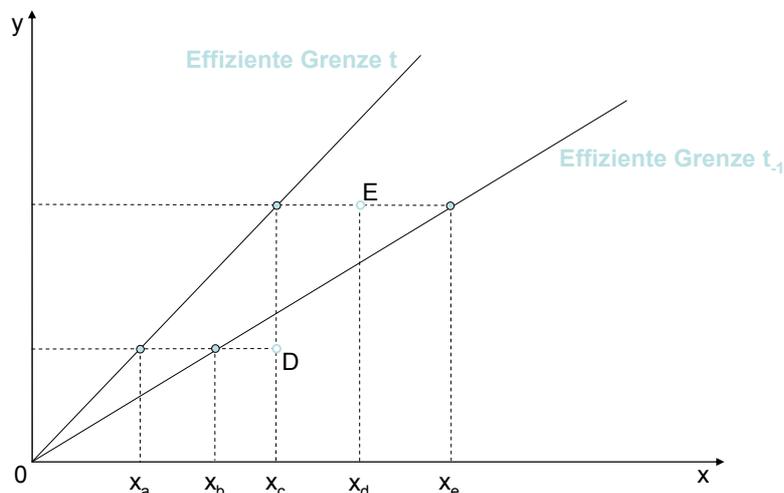
Betrachtet werden hier Malmquist Indices, die sich am effizienten Einsatz der Inputs orientieren, was in Wirtschaftssektoren, die sich durch monopolistische Strukturen auszeichnen, die Norm sein dürfte. Die Malmquist DEA misst die Veränderung der TFP zwischen zwei Produktionspunkten, die jeweils mit Hilfe von Distanzfunktionen nach Shephard zu einer Technologie in Bezug gesetzt werden.

³¹ Vgl. Kirjavainen und Loikkanen (1998).

³² vgl. Simar und Wilson (2000a und b).

³³ Malmquist Indices können sowohl mit Hilfe der DEA als auch auf Grundlage der SFA berechnet werden.

Abbildung 3.6: Input-orientierte Malmquist DEA (CRS)



wik

Definitionsgemäß nehmen Inputdistanzfunktionen einen Wert größer 1 an, wenn mehr Inputs eingesetzt werden, als im Vergleich zu einem effizienten Unternehmen notwendig sind, d.h. im Punkt D und bezogen auf die Technologie der ersten Periode gilt $x_c/x_b > 1$. In der zweiten Periode produziert das Unternehmen im Punkt E; ein einfaches Maß für die Variation der TFP zwischen zwei Perioden ist dann:³⁴

$$(3-8) \Delta TFP = TFP_1 / TFP_0 = \frac{x_c/x_a}{x_d/x_c}$$

In Abb. 3.6 nimmt Gleichung (3-8) einen Wert **größer 1** an, da gemessen in der Technologie der zweiten Periode (obere Gerade als gemeinsamer Bezugspunkt) der Abstand zur Grenze verkürzt werden konnte. Die Größe ist jedoch insofern zu unspezifisch, als sie die gesamte Veränderung des Produktionsergebnisses zwischen beiden Perioden einem Zuwachs an TFP zuschreibt, ohne Quellen genauer benennen zu können. Zudem wäre ΔTFP ebenso gut in Einheiten der ersten Periode messbar. Daher verwendet man häufig das geometrische Mittel der TFP-Indices zweier Perioden.

$$(3-9) \Delta TFP = TFP_1 / TFP_0 = \left[\frac{x_c/x_a}{x_d/x_c} \cdot \frac{x_c/x_b}{x_d/x_e} \right]^{0,5} = \underbrace{\frac{x_c/x_b}{x_d/x_c}}_{\Delta TE} \cdot \underbrace{\left[\frac{x_c/x_a}{x_c/x_b} \cdot \frac{x_d/x_c}{x_d/x_e} \right]^{0,5}}_{\Delta TF}$$

³⁴ Bei Inputorientierung ist darauf zu achten, dass die Produktivität reziprok zu den Distanzfunktionen ist. Daher erscheinen die Distanzen für TFP_0 im Zähler und für TFP_1 im Nenner.

Der Ausdruck nach dem Gleichheitszeichen in Formel 3-9 zerlegt nun die Variation der totalen Faktorproduktivität in ein Maß für die Verbesserung der technischen Effizienz zwischen zwei Perioden (Ausdruck vor der Klammer) und ein Maß des technischen Fortschritts – wiederum gemessen als geometrisches Mittel, d.h. $\Delta TFP = \Delta TE \cdot \Delta TF$. Dabei trifft Formel (3-9) explizit die Annahme einer Technologie mit konstanten Skalenerträgen; liegen diese nicht vor, so haben Färe et al. vorgeschlagen, den Index weiter zu zerlegen, indem zwischen CRS und VRS unterschieden wird.³⁵ In diesem Fall wird die Variation der TFP in drei Komponenten zerlegt:

$$(3-10) \Delta TFP = \Delta TE \cdot \Delta TF \cdot \Delta SE$$

Zwar ist diese Methodik, die TFP zu zerlegen, stark kritisiert worden, da der technische Fortschritt letztlich weiterhin gegen eine CRS-Technologie gemessen wird, wenngleich die Erweiterung von Färe et al. dazu gedacht war, Produktion unter VRS Bedingungen zu zulassen. Bisher hat sich jedoch keiner der alternativ gemachten Vorschläge weiter verbreiten können. Schließlich ist darauf hinzuweisen, dass Indices wie in Formel (3-9) nicht dazu geeignet sind, Fragen der allokativen Effizienz zu analysieren, da sie ohne explizite Preisinformationen ermittelt werden.

3.7 Anfälligkeit der DEA Technik für Datenfehler

Beachtet werden sollte, dass die DEA alle Abweichungen, die sich aus den Daten eines Unternehmens gegenüber der Frontier ergeben, als Ineffizienz bewertet. Im Gegensatz zu stochastischen Methoden, die über die Störterme der Schätzgleichungen zumindest einen Teil der Datenungenauigkeit auffangen können, sieht die DEA in ihrer Grundform eben keine Konfidenzintervalle oder ähnliches vor, weshalb fehlerhaft aufgenommene oder verarbeitete Daten entweder zu negativen Bewertungen einzelner Unternehmen führen oder aber die Lage der Grenze mitbestimmen können. Hiervon wären dann alle anderen Unternehmen negativ betroffen.

Es empfiehlt sich daher, die Daten, die im Rahmen der DEA Verwendung finden sollen, vor Durchführung der Berechnungen genau auf ihre Plausibilität zu kontrollieren.³⁶ Dies gilt auch und gerade, wenn diese Daten im Rahmen einer regulatorisch verfügbaren Erhebung von den betroffenen Unternehmen bezogen werden. Der Zeitaufwand bei einer DEA liegt insofern weniger in der eigentlichen Berechnung als darin die Daten auf- und vorzubereiten. Zu denken ist hier vor allem auch daran, dass alle Daten in denselben Einheiten oder Maßgrößen erfasst sein müssen, da ansonsten falsche

³⁵ Vgl. Färe et al. (1994).

³⁶ Auch bei stochastischen Methoden gehört ein Plausibilitätscheck der Daten zum „guten Ton“. Fehlerhafte Daten führen dort ebenfalls zu Verzerrungen der Ergebnisse. Es sollte mit der gleichen Sorgfalt an die Datenaufbereitung wie bei der DEA herangegangen werden. Daher gelten die folgenden Ausführungen zum Aufwand der Datenaufbereitung im wesentlichen unverändert auch für die ökonomischen Verfahren. Allerdings wiegen einzelne Datenfehler nicht ganz so schwer wie bei DEA, die sozusagen „no margin of error“ zulässt.

Ergebnisse ermittelt werden. Es ist in diesem Zusammenhang häufig hilfreich, die Werte bezüglich einzelner Inputs und Outputs in Diagrammen wie Abb. 3.5 abzutragen, um so bereits rein optisch auf Ausreißer schließen zu können. Jedoch sei nochmals darauf hingewiesen, dass dieses Problem ähnlich wie die Schlupf-Problematik vor allem durch den Umfang der Untersuchung getrieben wird – je mehr Unternehmen verglichen werden, desto weniger wird das Gesamtergebnis durch ein falsches Datum beeinflusst. Zudem sind verfälschte Ergebnisse aus der DEA vor allem dann gefährlich, wenn diese mechanisch in regulierungsrelevante Größen (wie z.B. X-Faktoren) umgesetzt werden. Hierauf war bereits an anderer Stelle dieser Darstellung hingewiesen worden.

Die Anfälligkeit der DEA für falsche Daten lässt sich zudem reduzieren, wenn zusätzlich zu den hier dargestellten Grundmodellen auch eine sog. Chance Constrained DEA betrachtet wird.³⁷ Derartige abgewandelte LP sehen vor, dass beide Seiten des Produktionsprozesses durch stochastisches Verhalten geprägt sein können, d.h. die Analyse bezieht sich auf Mittelwerte und sowohl Inputs als auch Outputs müssen **nicht zwingend auf einer** Seite der analysierten effizienten Grenze liegen. Vielmehr werden die Restriktionen der Optimierung so gewählt, dass zumindest ein Teil der Beobachtung auch **außerhalb** der Grenze liegen kann.³⁸ Aus Sicht der Regulierungstheorie ist ein solches Vorgehen ein Weg, regulatorische Diskretion zu formalisieren, da im Endeffekt die Chance Constrained DEA oder die vorsichtige Anwendung von Ergebnissen aus den Grundmodellen zu ähnlichen Ergebnissen – weicheren Effizienzbewertungen – führen dürften. Diese „Weichheit“ der Bewertung wird insofern von der DEA auch dadurch unterstützt, dass sie so konstruiert ist, die Unternehmen in ein möglichst günstiges Licht zu stellen, da allen Beobachtungen im Zuge der linearen Programmierung ein möglichst großer Wert zugewiesen wird (vgl. Annex 2).

3.8 DEA und Energieversorgungsunternehmen

DEA ist eines der am häufigsten genutzten Instrumente zum Benchmarking von Verteilunternehmen in der Elektrizitätswirtschaft. Während Deutschland erst am Beginn der Diskussionen zur Einführung von Benchmarking Methoden steht, sind die Vorbereitungen in vielen anderen europäischen (u.a. Österreich, Niederlande, Schweden, Norwegen, Finnland) und außereuropäischen Staaten (u.a. Australien, Vereinigte Staaten) bereits sehr weit gediehen bzw. wird diese Methodik seit einiger Zeit angewendet. In einigen Ländern kommen dabei mehrere Verfahren komplementär

³⁷ Vgl. Olesen (2004).

³⁸ Damit werden formal Standardabweichungen und Konfidenzintervalle für den Produktionsprozess formuliert. Für den Fall mehrerer Inputs oder Outputs sind dann auch Kovarianzen zu analysieren, wobei häufig die Annahme getroffen wird, dass zwischen den Abweichungen der einzelnen Komponenten keine Beziehungen bestehen.

zur Anwendung.³⁹ So werden im Vereinigten Königreich beispielsweise DEA und Regressionsanalysen parallel verwendet. Der australische Bundesstaat New South Wales setzt hingegen auf die Kombination aus DEA und SFA. Allerdings haben nicht alle Staaten positive Erfahrungen mit diesem Instrumentarium gemacht, namentlich die Niederlande und Österreich, die beide ihre Vorhaben, DEA zur Preisregulierung einzuführen, vorerst zurückziehen mussten. Die Erfahrung dort hat gezeigt, dass, wenn bei geringen Erfahrungen mit diesem Instrument sowohl auf Seiten der Regulierungsbehörde als auch auf Seiten der Netzbetreiber Preisvorgaben festgesetzt wurden, es zu harten und langwierigen rechtlichen Auseinandersetzungen zwischen beiden Gruppen kommen kann. In beiden Fällen wurden über einen langen Zeitraum personelle Ressourcen gebunden, um (vermeintliche) Fehlansätze der Regulierung im Nachhinein zu korrigieren. Während das Instrument in Österreich aufgrund der fehlenden Akzeptanz auf Seiten der Regulierten erst gar nicht zur Anwendung kam, führte die erstmalige Durchführung einer DEA in den Niederlanden zu heftigen Protesten, die sich vor allem an der mechanischen Umsetzung der DEA-Ergebnisse in Vorgaben für den X-Faktor festmachten. Um dies in Deutschland zu vermeiden, wäre also ein Einführungsprozess unter Einbindung der betroffenen Unternehmen anzuraten, der Erfahrungswerte schafft, die Informationslage auf beiden Seiten verbessert und somit den Boden für zukünftige Akzeptanz bereitet.

Viel Diskussionen gibt es bei der Einführung derartiger Verfahren regelmäßig auch über die anzuwendenden Parameter. Tabelle 3.1 zeigt einen Überblick über die Häufigkeit des Einsatzes wichtiger Input- und Outputvariablen in ausgewählten Benchmarking Studien zu Energieversorgungsunternehmen, um einen Eindruck von Variablen einer möglichen zukünftigen Datenbasis zu geben. Eine eindeutige Tendenz ist jedoch nicht ersichtlich. Für aussagekräftige Schlussfolgerungen über die Art und Anzahl der zu berücksichtigenden Variablen wären umfangreiche Voruntersuchungen anhand konkreter Datensätze erforderlich, um die Bedeutung bestimmter Parameter im Falle der deutschen Netzbetreiber fundiert ab- und einschätzen zu können.

³⁹ Zu Übersichten über den Stand des Regulierungsprozesses sowie die verwendeten Benchmarking Verfahren siehe z.B. Hense und Schäffner (2004), Jamasb und Pollitt (2001, 2003) sowie Ajodhia et al. (2003).

Tabelle 3.1: Häufigkeit des Einsatzes wichtiger Input- und Outputvariablen in 20 Benchmarking Studien zu Energieversorgungsunternehmen

Inputvariable (Anzahl verwendeter Instanzen)	Outputvariable (Anzahl verwendeter Instanzen)
Verkaufte Einheiten Gesamt (2)	Gesamt (12); Wohnviertel (6); Nicht-Wohnviertel (6)
Kundenzahlen Gesamt (1)	Gesamt (11); Wohnviertel (5); Nicht-Wohnviertel (5)
Größe des Netzwerkes Größe des Netzwerkes (11); Niederspannung (2); Mittelspannung (1), Hochspannung (2)	Größe des Netzwerkes (4)
Kapazität der Transformatoren Gesamt (11); Mittelspannung (1); Hochspannung (1)	Gesamt (1); Anzahl der Transformatoren (1)
Versorgungsgebiet Versorgungsgebiet (2)	Versorgungsgebiet (6)
Höchstleistung Höchstleistung (1)	Höchstleistung (4)
Eingekaufte Energie Eingekaufte Energie (2)	An andere Energieversorgungsbetriebe verkaufte Energie (1)
Verluste Übertragungs- / Verteilungsverluste (4)	---
Arbeit Arbeit/ Löhne (15); Verwaltungsarbeit (1); Technische Arbeit (1)	---
Kostenmessung Betriebsaufwand (7); Betriebsaufwand + jährliche standardisierte Kapitalkosten (1); Verwaltungs- /Buchführungskosten (2); Wartungskosten (1); Kapital (5); Investitionsaufwandskosten + Arbeitskosten (1); Materialien (1)	---
Verschiedenes Industrielle Nachfrage (1); Kundenstreuung (2); Anteil industrieller Energie (1); Größe des Netzwerkes/ Kunden (1); % der Systemlastung (1); Wohnbezirks-/ Gesamtverkäufe (1); Stromausfall (1); Wohnviertel- kunden / Netzwerkgröße (1); Vorräte (1); Kabellänge x Spannung (1)	Service-Zuverlässigkeit (1); Auslastungsfaktor (1); Netto-Gewinnspanne (1); Einkünfte (1); Entfernungs-index (1); Netzwerkdichte (1); Kategorische Variable für Stadtgebiete (1)

Quelle: Jamasb und Pollitt 2003

4 Stochastische Verfahren

Grundlage aller stochastischen Verfahren ist die Annahme, dass die zu beurteilenden Datensätze aus einer gemeinsamen Verteilung gezogen werden. In Bezug auf die hier zu analysierenden Stromverteilunternehmen würde diese Annahme daher besagen, dass die Unternehmen mit einer Technologie arbeiten, die über alle Unternehmen hinreichend ähnlich ist, um aus der Gesamtheit der Unternehmen – ihrer Kosten- oder Produktionsstruktur – auf die von ihnen verwendete Technologie schließen zu können. In der praktischen Anwendung bedeutet das Verwenden stochastischer Schätzgleichungen, dass ex ante festzulegen ist, welche Annahmen über den Verlauf der Produktionsfunktion gelten sollen.

Im Gegensatz zur DEA werden damit wesentlich restriktivere Annahmen getroffen, auf deren Grundlage die Daten untersucht werden. Während sich DEA Ergebnisse jedoch nur schwer überprüfen lassen, da eben keine Verteilungsannahmen bezüglich der ermittelten Effizienzwerte vorliegen, ist dies bei statistischen Verfahren möglich. Die a priori getroffenen Annahmen über eine eigentlich unbekannt Technologie lassen sich mit Hilfe statistischer Testverfahren auf ihre Plausibilität hin überprüfen, d.h. ob es sich um eine zulässige Beschreibung der Daten handelt oder nicht.

4.1 Alternative Spezifikationen der Produktionsfunktion

Grundsätzlich bieten stochastische Ansätze eine Vielzahl von Möglichkeiten, denn neben Produktionsfunktionen – von denen in diesem Text bisher zumeist ausgegangen wurde – erlauben Regressionsverfahren unter bestimmten mikroökonomischen Voraussetzungen auch Profit- oder Kostenfunktionen zu schätzen. Ein solches Vorgehen beinhaltet dann die Prämisse, dass die verwendete Technologie ausreichend prägend für das Unternehmen ist, um ihre Form quasi indirekt aus Kosten- oder Profitverläufen zu schließen.⁴⁰ Dabei sollte die Annahme über die Art der Produktion/Technologie flexibel genug sein, um mehr als eine mögliche Produktionsstruktur abbilden zu können. Andererseits muss sie gerade bei Kosten- oder Profitfunktionen einer Reihe von Anforderungen genügen, um den Rückschluss auf die Produktionsfunktion zu erlauben: So muss eine Kostenfunktion z.B. homogen

⁴⁰ So gilt z.B. für Produktionsfunktionen vom Cobb-Douglas-Typ aufgrund der Dualität, dass aus dem Verlauf der Produktionsfunktion direkt auf die Kostenfunktion geschlossen werden kann und umgekehrt. Insofern zeigt die Produktionsfunktion also fallende Skalenerträge. Wegen Shephard's und Hotelling's Lemma lässt sich zudem direkt aus den Kosten- oder Profitfunktionen auf den Verlauf der Inputnachfragefunktionen schließen. Diese Eigenschaft der Cobb-Douglas-Funktion lässt sich auch durch ihre konstante Substitutionselastizität von 1 beschreiben, die dazu führt, dass Änderungen des Faktorpreisverhältnisses sich c.p. direkt in Änderungen der Faktornachfrage übertragen.

vom Grade 1 in allen Preisen sein, d.h. eine Verdopplung der Inputpreise muss c.p. auch eine Verdopplung der Kosten nach sich ziehen.⁴¹

Betrachtet werden zunächst lediglich Produktionsfunktionen. Die Literatur kennt eine ganze Reihe möglicher Formen. In der ökonometrischen Analyse spielt jedoch nur ein Teil dieser Funktionen eine Rolle. Aufgrund ihrer einfachen Handhabung und Interpretation wurden Effizienzschätzungen in der Vergangenheit häufig mit der Cobb-Douglas-Produktionsfunktion vorgenommen: ⁴²

$$(4-1) y = Ax_1^{\beta_1} x_2^{\beta_2} \cdot \exp\{v\}$$

wobei es sich häufig anbietet, nicht die Produktionsfunktion selbst sondern ihre logarithmierte Form zu schätzen:

$$(4-2) \ln y = \ln A + \beta_1 \ln x_1 + \beta_2 \ln x_2 + v$$

Gleichung (4-2) ist linear und insofern relativ einfach durch die Kleinstquadratmethode, die eine lineare Beziehung zwischen der zu erklärenden und den erklärenden Variablen unterstellt, zu schätzen. Hinzu kommt, dass die Regressionskoeffizienten die Produktionselastizitäten der Cobb-Douglas-Funktion bilden und als solche direkt interpretierbar sind. Nachteilig wirkt dagegen, dass Gleichung (4-2) konstante Skalenerträge hat, die wiederum durch die Summe der Regressionskoeffizienten gegeben sind und die Substitutionselastizität der Funktion immer 1 beträgt. Die Folge dieser Restriktionen ist, dass deutliche Unterschiede in der Unternehmensgröße bei Schätzungen mit der Cobb-Douglas nicht analysiert werden können. Auch ist nicht zulässig, dass die Skalenerträge sich verändern, während die Produktion ausgeweitet wird.

Eine flexiblere Form stellt die sog. Translog⁴³ dar, bei der sich jedoch die Zahl der zu schätzenden Regressionsparameter deutlich erhöht, was im Fall weniger Beobachtungen problematisch sein kann, da u.U. weniger Datensätze als Koeffizienten vorhanden sind; also Probleme bezüglich der Freiheitsgrade auftreten. ^{44 / 45}

⁴¹ Für eine ausführliche Diskussion dieser Fragen sei auf Kumbhakar und Lovell (2003) oder Mas-Collel et al. (2004) verwiesen. Bezüglich einer Kostenfunktion werden aber im allgemeinen fünf Annahmen getroffen: 1) Wie bereits erwähnt, ist die Kostenfunktion homogen vom Grade 1 in allen Preisen. 2) Die Kosten sind größer oder gleich Null für alle Inputpreise größer oder gleich Null und für positive Outputmengen. 3) Haben zwei Produktionssituationen unterschiedlich hohe Kosten und unterscheiden sich nur bezüglich eines Inputpreises, so muss dieser Preis im Falle höherer (niedrigerer) Kosten ebenfalls höher (niedriger) sein. 4) Die erste Ableitung der Kostengleichung nach den Preisen enthält diese nicht mehr als Argument. 5) Die Kostenfunktion ist schwach konkav in den Inputpreisen, wenn die Produktionsfunktion strikt quasi konkav ist.

⁴² Die Formulierung $\exp\{v\}=e^v$ für den Störterm wird hier gewählt, um in der logarithmierten Form wieder eine additive Verknüpfung zu erreichen, ohne dass eine Konstante mitgeführt werden muss.

⁴³ Eigentlich „transcendental logarithmic production function“ – letztlich ein Taylor-Polynom zweiten Grades, das eine höchst flexible Anpassung der Produktionsfunktion an die Daten ermöglicht.

⁴⁴ Um die Anzahl der zu schätzenden Koeffizienten zu beschränken, wird daher im Allgemeinen Symmetrie hinsichtlich der β_{ij} unterstellt, d.h. $\beta_{ij}=\beta_{ji}$.

$$(4-3) \ln y = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \frac{1}{2} [\beta_{11} (\ln x_1)^2 + \beta_{22} (\ln x_2)^2] + \beta_{12} \ln x_1 x_2 + v$$

Zusätzlich kann nicht ausgeschlossen werden, dass die Regressionsparameter von Funktionen wie (4-3) sich als multikollinear erweisen, d.h. zwischen den einzelnen Regressionsvariablen kann eine lineare Beziehung dergestalt bestehen, dass eine oder mehrere Linearkombinationen der Variablen existieren, die sich zu Null addieren.⁴⁶ Zu bemerken ist, dass sich die Wahrscheinlichkeit, mit der das Problem auftritt, erhöht, wenn die zu analysierenden Datensätze klein im Verhältnis zur Menge der zu schätzenden Variablen sind.

Gerade wenn längere Datenreihen vorliegen, ist es zudem interessant, den technischen Fortschritt explizit zu modellieren: Hier gibt es seitens der ökonomischen Theorie mehrere Methoden der Darstellung, mit denen sich aber auch jeweils inhaltliche Aussagen verbinden. Modelliert man den technischen Fortschritt wie in Gleichung (4-2'), so verbindet sich damit die Annahme, dass es sich um Hicks-neutralen technischen Fortschritt handelt.

$$(4-2') \ln y = \ln A + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \beta_3 t + v \quad \forall t \geq 1$$

In diesem Fall verbessert sich im Zeitablauf zwar die Ausbeute des Produktionsprozesses und damit die TFP, die optimale Faktoreinsatzkombination bleibt hiervon aber unberührt. Dagegen tritt in Gleichung (4-3') technischer Fortschritt auf, der alle Faktoren beeinflusst.

$$(4-3') \ln y = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \beta_t t + \frac{1}{2} [\beta_{11} (\ln x_1)^2 + \beta_{22} (\ln x_2)^2] + \beta_{12} \ln x_1 x_2 + \beta_{1t} \ln x_1 t + \beta_{2t} \ln x_2 t + \beta_{tt} t^2 + v$$

Dem Anwender sind daher eine Vielzahl von möglichen Untersuchungsstrukturen an die Hand gegeben, deren Einsatzmöglichkeiten nicht zuletzt vom Datenbestand bestimmt werden. Insbesondere bei umfangreichen Datensätzen bietet es sich aber an, die Translog-Funktion zu verwenden, da dann a priori keine Annahme über die Struktur des Produktionsprozesses getroffen werden muss.

45 Eine grobe Daumenregel in der Statistik besagt, dass ab einer Stichprobengröße von ca. 50 Beobachtungen davon ausgegangen werden kann, eine erste Annäherung an eine Normalverteilung zu erreichen, wobei kleinere Umfänge als kritisch anzusehen sind. Bei einer Anzahl von über 900 Netzbetreibern in Deutschland – so Daten für alle verfügbar sind – kann von einem mehr als hinreichend großen Stichprobenumfang ausgegangen werden. Die Zahl der aufzunehmenden exogenen Variablen lässt sich ferner über das korrigierte Bestimmtheitsmaß kontrollieren. Steigt dieses durch eine zusätzlich aufgenommene Variable, verbessert sich auch der Erklärungsgehalt des ökonometrischen Modells. Im Gegensatz dazu lässt das einfache Bestimmtheitsmaß R^2 keine Aussage zu, da dessen Wert sich mit jeder zusätzlichen Variable erhöht – unabhängig vom Erklärungsgehalt. Bei 900 Beobachtungen im Querschnitt kann davon ausgegangen werden, dass Schätzungen mit 15 bis 20 Parametern unproblematisch sein sollten.

46 Ganz offensichtlich wird diese Möglichkeit bereits in der Formulierung der Translog zumindest in Kauf genommen, denn mehr als 50 % aller Variablen der Funktion (4-3) sind Linearkombinationen im obigen Sinne. Ein erster Hinweis, dass Multikollinearität vorliegen könnte, besteht in hohen Werten für das Bestimmtheitsmaß R^2 , die mit insignifikanten T-Werten für die einzelnen Regressionsvariablen einhergehen.

4.2 Traditionelle Ansätze zur Schätzung von Effizienzgrenzen

Um eine Schätzung durchzuführen, muss aber nicht nur eine Entscheidung bezüglich der Annahmen über den tatsächlichen Produktionsprozess und damit bezüglich der verwendeten Schätzgleichung getroffen werden; auch bezüglich der zu verwendenden Schätzmethode sieht sich der Anwender Freiheitsgeraden gegenüber. Wie im folgenden zu sehen sein wird, sind viele der im Bereich der Effizienzanalyse eingesetzten Methoden letztlich Weiterentwicklungen bzw. Abwandlungen der Methode der kleinsten Quadrate, die als eine Art Ausgangspunkt gelten kann.

4.2.1 Methode der kleinsten Quadrate (OLS)

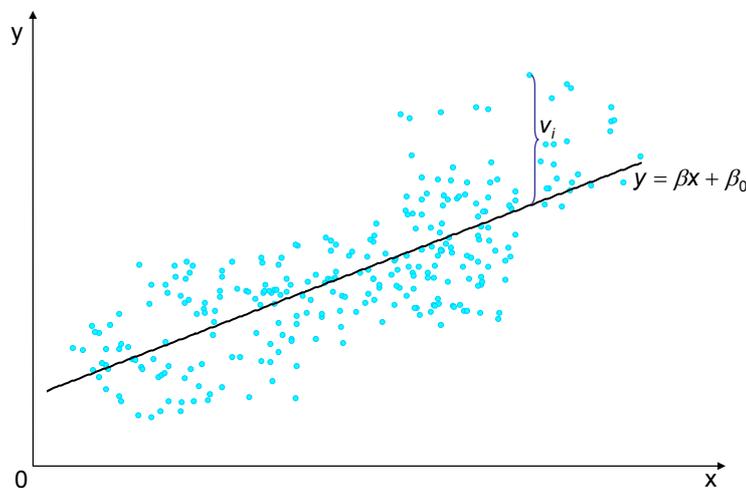
Die Methode der kleinsten Quadrate (ordinary least squares, OLS) dient dazu, im Rahmen einer Ausgleichsrechnung den Wert von zu schätzenden Parametern (d.h. den Regressionsvariablen) so zu bestimmen, dass die Summe der quadrierten Abweichungen minimiert wird.⁴⁷ Eine weit verbreitete Anwendung stellt die lineare Regressionsanalyse dar.

Wie Abb. 4.1 zu entnehmen ist, gibt die (lineare) Schätzgleichung letztlich einen **Erwartungswert** für das durchschnittliche Verhalten aller Punkte/Unternehmen in der untersuchten Datengruppen wieder – die Schätzgerade verläuft durch die Mitte der Punktwolke, was aus dem Ansatz der Kleinstquadratschätzungen folgt, der positive und negative Abweichungen auszugleichen sucht. Die Abweichungen einzelner Unternehmen von den im Mittel erwarteten Werten werden in den Residuen v_i zusammengefasst, deren Summe einen Erwartungswert von Null hat, da sich im Mittel wiederum die positiven und negativen Abweichungen gerade aufheben.

Die lineare Schätzgleichungen in Abb. 4.1 beschreibt die Beziehung zwischen einem Input x und einem Output y , wobei sich die Analyse problemlos auf mehrere Inputs erweitern ließe. Problematischer ist es dagegen, wenn mit einer derartigen Schätzgleichung Unternehmen untersucht werden sollen, die mehrere Produkte erzeugen. Im Allgemeinen müssen dann die einzelnen Outputs in geeigneter Weise zu einer Gesamterzeugungsmenge aggregiert werden. Hierbei stellen sich dann wieder die Fragen der Gewichtung der einzelnen Outputs zueinander, wie sie bereits im Abschnitt zu den PIN thematisiert worden waren.

⁴⁷ Derart ermittelte Werte für die Regressionskoeffizienten sind sog. beste, lineare, erwartungstreue Schätzer, d.h. es lässt sich beweisen, dass es keinen linearen erwartungstreuen Schätzer gibt, dessen Varianz kleiner ist. In der englischsprachigen Literatur ist es üblich, solche Schätzer BLUE zu nennen (BLUE = best linear unbiased estimator).

Abbildung 4.1: Methode der kleinsten Quadrate



Aus Sicht einer Analyse, die damit befasst ist, die Effizienz einzelner Unternehmen oder eines ganzen Sektors zu untersuchen, besteht das Problem nun darin, dass die OLS-Schätzgleichung von ihrer Konzeption her eben gerade nicht zum Ziel hat, die bestmögliche Performance abzubilden. Im Gegenteil vermischen sich aus Sicht einer an der Effizienz orientierten Regulierungsaufgabe best und worst practice. Folge einer solchen Durchschnittsbetrachtung muss sein, dass etwa 50% der Unternehmen mittelfristig gezwungen sind, sich dem Durchschnitt anzunähern. Dagegen bleibt die andere Hälfte der Unternehmen unbehelligt, obwohl die Analyse keinen Hinweis darauf geliefert hat, dass es sich bei dieser zweiten Hälfte ausschließlich um effiziente Unternehmen handelt. Die hier verborgenen Potenziale können also nicht wohlfahrtssteigernd genutzt werden. Die dynamische Wirkung eines solchen Vorgehens ist zudem mit Zweifeln behaftet, da nicht ausgeschlossen werden kann, dass diejenigen Unternehmen, die besser als der Durchschnitt liegen, durch einen an einer OLS-Schätzung geleiteten Regulierungsansatz mit nachteiligen Anreizen konfrontiert sind. Die regulatorische Unsicherheit, nicht zu wissen, wie man sich relativ zum Sektor stellt, schwindet. Es ist anzunehmen, dass sich Unternehmen als Reaktion auf die regulatorische „Bestätigung“ ihrer relativen Güte in der Folge weniger effizienzorientiert zeigen. Dann aber dürfte die Punktwolke im Zeitablauf enger zusammenrücken, d.h. die v_i würden kleiner, jedoch würde sich die Regressionsgerade nicht bewegen. Im Gegensatz zur Gesamtwirtschaft würde die Produktivitätsentwicklung im Sektor also stagnieren.

4.2.2 Corrected Ordinary Least Squares (COLS)

Den obigen Problemen einer Durchschnittsbetrachtung infolge des Einsatzes von OLS kann auf mehrere Arten und Weisen begegnet werden. So lässt sich die Schätzung einer effizienten Grenze auf Grundlage der kleinsten Quadrate in einem zweistufigen Verfahren bewältigen:

- Zunächst wird eine einfache OLS-Schätzung durchgeführt, von deren Ergebnis bezüglich der Steigungsparameter β_i angenommen wird, dass es auch im Hinblick auf die effiziente Grenze Gültigkeit hat. D.h. auch für das Grenzunternehmen kann nur angegeben werden, wie sich seine Produktion oder seine Kosten durchschnittlich entwickeln. Damit wird praktisch unterstellt, dass das effiziente Unternehmen mit einer ähnlichen Technologie arbeitet wie alle anderen Unternehmen des Sektors und lediglich bessere TFP-Werte erreicht.
- In einem zweiten Schritt wird dann der aus Sicht einer Frontier-Analyse verzerrte Schätzer des Achsenabschnitts der Regressionsfunktion β_0 korrigiert. Hierbei spielt die Frage der grundsätzlich unterstellten Verhaltensannahme (Kostenminimierung vs. Produktionsmaximierung) eine entscheidende Rolle, denn sie ist entscheidend für die Richtung der erforderlichen Korrektur. Bei Kostenminimierung wird die Regressionsgerade nach unten korrigiert, bei Produktionsmaximierung hingegen nach oben verschoben.

Bezogen auf eine Produktionsschätzung, wie sie Abb. 4-2 zeigt, wird der Schätzer des Achsenabschnitts dann so ermittelt, dass

$$(4-4) \beta_0^* = \beta_0^{OLS} + \max_i(v_i^{OLS})$$

gilt. In der gleichen Art und Weise werden auch alle anderen Residualterme korrigiert:

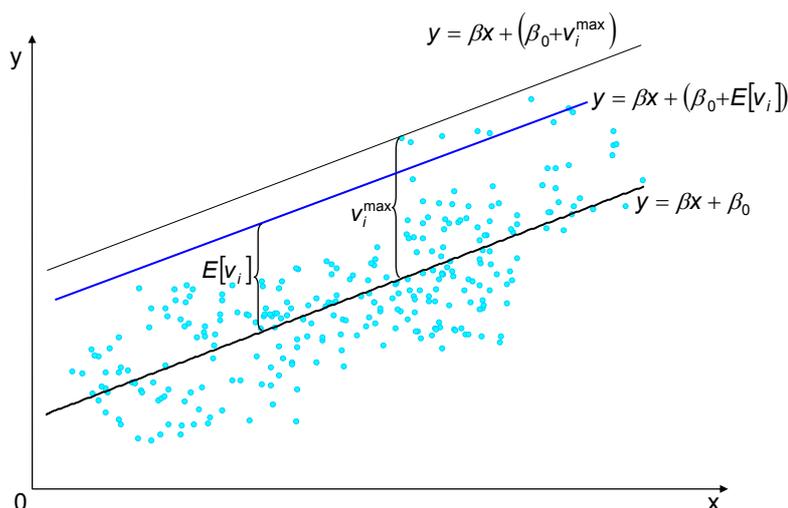
$$(4-5) -v_i^* = v_i^{OLS} - \max_i(v_i^{OLS})$$

Dadurch kann die technische Effizienz eines einzelnen Produzenten ermittelt werden, wobei im Falle einer logarithmierten Schätzfunktion wieder $TE_i = \exp[-v_i^*]$ gilt.

Geht man von normalverteilten v_i aus, d.h. von einem arithmetischen Mittel von Null, so resultiert aus diesem Vorgehen eine neue Normalverteilung, deren Mittelwert dann $-\max_i(v_i^{OLS})$ beträgt. Damit werden die gesamten Störterme als Ineffizienz beschrieben. Dagegen findet keine Beachtung, dass sich auch aus von den Unternehmen nicht zu verantwortenden Umständen Produktionsverschiebungen

ergeben können. Diese werden von der COLS in vollem Umfang als Ineffizienz interpretiert – es handelt sich insofern eigentlich nicht um ein stochastisches sondern um ein deterministisches Modell.

Abbildung 4.2: Korrektur einer Produktionsschätzung bei COLS und MOLS



Die COLS Technik ist einfach anzuwenden und erzeugt eine effiziente Grenze, die für den effizienten Produzenten über den Daten aller anderen Unternehmen liegt. Wie die Abbildung zeigt, verläuft die COLS Schätzung (obere Linie) parallel zur ursprünglichen OLS Gerade (untere Linie). Wie bereits angedeutet wurde, wird damit implizit angenommen, dass die effiziente Technologie dieselbe Struktur wie die durchschnittliche Technologie aufweist. Diese Annahme wird in Teilen der Literatur als wesentlich zu restriktiv betrachtet, da anzunehmen ist, dass die Effizienz bestimmter Produzenten sich auch in überlegener Technologie ausdrückt.

Unserer Ansicht nach ist COLS für die praktische Anwendung daher deutlich zu restriktiv. Hinzu kommt, dass das Verfahren aus praktisch regulatorischer Sicht deutliche Risiken birgt, da Entscheidungen letztlich durch einen einzelnen Datenpunkt motiviert werden. Ist dieser Datenpunkt aber ein Ausreißer, was gerade bei Extrempunkten nicht unbedingt auszuschließen ist, so werden die aus der COLS abgeleiteten Effizienzziele die Unternehmen u.U. deutlich überfordern. Ähnliche Effekte würden sich ergeben, wenn ein Unternehmen in einer bestimmten Periode außergewöhnliche Erfolge erzielen konnte, die es in der Folge nicht wiederholen kann – hier wird deutlich, dass die einseitige Interpretation der Störterme als relative Ineffizienz auch für das beste Unternehmen nicht ungefährlich ist.

4.2.3 Modified Ordinary Least Squares (MOLS)

Bezüglich des Ansatzes der MOLS ist zunächst festzustellen, dass der Begriff in der Literatur für eine Reihe von verwandten Ansätzen gebraucht wird. Einigkeit besteht bezüglich der Interpretation von MOLS als eine Methode, die versucht, die Ergebnisse der einfachen Korrektur-Ansätze zu verallgemeinern oder weiterzuentwickeln. Der entscheidende Unterschied zwischen beiden Ansätzen besteht nun darin, dass die MOLS-Ansätze eine explizite Annahme bezüglich der Gestalt der Störterme treffen. Dabei wird angenommen, dass die Residuen weiterhin Ausdruck technischer Ineffizienz sind, wobei größere Abweichungen von der effizienten Grenze mit geringeren Wahrscheinlichkeiten belegt werden. Meist werden MOLS unter der Annahme einer Halbnormalverteilung oder einer exponentiellen Verteilung geschätzt.

So verhält sich z.B. die Wahrscheinlichkeitsdichtefunktion einer exponential verteilten Zufallsvariablen v entsprechend Gleichung (4-6): ⁴⁸

$$(4-6) f(v) = \frac{1}{\sigma_v} \cdot \exp\left\{-\frac{v}{\sigma_v}\right\}$$

Es gibt vor allem drei Unterschiede zwischen einer solchen Schätzung der Produktionsfunktion und der Analyse einer Produktionsgrenze mit der Standardnormalverteilung: Halbnormal oder exponential verteilte Variablen nehmen nur Werte größer oder kleiner Null an, während ihr Modalwert weiterhin bei Null liegt und ihre Standardabweichung von 1 verschieden sein kann. Als Folge dieser Annahme fallen Modalwert und arithmetisches Mittel von v auseinander. Die Korrektur des Achsenabschnitts und die Ermittlung der einzelnen Effizienzwerte erfolgt dann in ähnlicher Weise, wie bereits in der Diskussion der COLS dargestellt:

$$(4-7) \beta_0^* = \beta_0^{OLS} + E(v_i) \text{ mit } E(v_i) \geq 0$$

$$(4-8) -v_i^* = v_i^{OLS} - E(v_i)$$

Auch die MOLS-Methode kann relativ einfach angewandt werden, sie beinhaltet jedoch das Risiko, dass nicht alle Unternehmen von der ermittelten Grenze eingeschlossen werden (angedeutet durch die mittlere Linie in Abbildung 4-2, wo einige Beobachtungen über der Geraden liegen). Es kann nämlich nicht ausgeschlossen werden, dass für einzelne Beobachtungen $[v_i^{OLS} - E(v_i)] > 0$ gilt – es also Effizienzbewertungen gibt, die oberhalb von 1 liegen. Möglich ist auch, dass gar kein Unternehmen die effiziente Grenze erreicht, die wiederum durch die Parameter der OLS Schätzung geprägt ist, weshalb alle gegen die COLS in diesem Zusammenhang geäußerten Kritikpunkte auch für die MOLS gelten müssen. Vor allem die Frage des Nichtunterscheidens zwischen

⁴⁸ Praktisch werden bereits die MOLS-Ansätze zumeist mit der Maximum-Likelihood Methode, einer asymptotischen Annäherung an den wahren Wert, gelöst (vgl. Abschnitt 4.3.1.1).

zufälligen Ereignissen und Ineffizienzen bleibt problematisch. Hinzu kommt, dass die Verteilungsannahmen über die v_i unbegründet bleiben, weshalb unserer Ansicht nach auch die MOLS aus regulatorischer Sicht einen zu großen Unsicherheitsfaktor darstellen.

4.3 Stochastic Frontier Analysis (SFA)

An den COLS und MOLS Ansätzen war vor allem kritisiert worden, dass sie nicht zwischen Ineffizienzen und zufälligen Schocks unterscheiden. An diesem Punkt setzen die SFA-Ansätze an und ergänzen den weiterhin symmetrisch verteilten Störterm v_i , der Messfehler und andere Zufälle (Streiks, Unwetter etc.) einfängt, um eine weitere nicht negative Zufallsvariable u_i , die Ausdruck von Ineffizienzen ist, so dass im Cobb Douglas Fall für ein einzelnes Unternehmen die stochastische Produktionsgrenze in Gleichung (4-9) gilt.

$$(4-9) \ln y_i = \ln A + \beta_{1i} \ln x_{1i} + \beta_{2i} \ln x_{2i} + \varepsilon_i \quad \forall \varepsilon_i = v_i - u_i$$

Damit hängt die relative Bewertung der Performance eines Unternehmens nunmehr von zwei Größen ab: Einerseits wird ein Unternehmen die Produktionsgrenze nicht erreichen, wenn beide Störterme negativ wirken, wobei allerdings nur die u_i zur Ermittlung der relativen Effizienz herangezogen werden. Andererseits ist es durchaus möglich, dass ein Unternehmen oberhalb der Frontier liegt, wenn der reine Noise-Term v_i einen stark positiven Wert annimmt.⁴⁹ Dabei wird regelmäßig angenommen, dass die beiden Störterme unabhängig voneinander und von den Regressionskoeffizienten verteilt sind. Grundsätzlich beruht die Anwendung von SFA dabei immer auf der zunächst unbewiesenen Hypothese, dass die Unternehmen der Grundgesamtheit sich durch Ineffizienzen auszeichnen. Ist dies der Fall, sollte der Störterm eine Schiefe hin zu negativen Werten aufweisen, andernfalls jedoch normal verteilt sein, sodass eine OLS Schätzung resultiert. Bezüglich dieser Frage hat Coelli einen Test vorgeschlagen:⁵⁰

$$(4-10) \text{ Schiefema\ss} = \frac{m_3}{\sqrt{\frac{6m_2^3}{N}}}$$

mit

m_2 : Moment 2. Ordnung der Wahrscheinlichkeitsverteilung (Varianz);

⁴⁹ Den v_i kommt hier die Aufgabe zu, Schätzfehler bei der Bestimmung der effizienten Grenze aufzufangen und so zu verhindern, dass Ausreißer zu großen Einfluss auf die Bestimmung ihrer Lage nehmen.

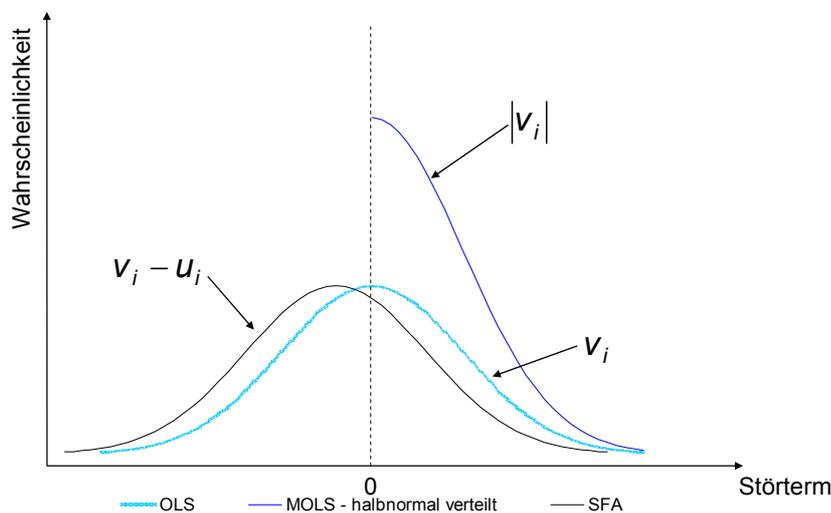
⁵⁰ Vgl. Coelli (1995).

m_3 : Moment 3. Ordnung;

N : Anzahl der Beobachtungen.

Die gemeinsame Dichtefunktion von u und v ergibt sich aus dem Produkt der individuellen Dichtefunktionen – ein Verlauf wie in Abb. 4.3 wird sich daher nur ergeben, wenn sich beide Störterme lediglich in Bezug auf ihren Mittelwert unterscheiden, ihre Standardabweichungen aber gleich sind.

Abbildung 4.3: Mögliche Verteilungen der Produktionseffizienz bei stochastischen Verfahren



4.3.1.1 Maximum Likelihood Schätzung und SFA

Im Gegensatz zu den bisher diskutierten Verfahren ist es üblich, die Parameter von stochastischen Grenzen durch sog. Maximum Likelihood Schätzer zu bestimmen. Der wichtigste Grund hierfür ist, dass bei diesem Verfahren die effizienten Unternehmen bei der Ermittlung der Regressionskoeffizienten eine größere Rolle spielen, als dies bei einer letztlich OLS basierten Analyse der Fall ist. Hinzu kommt, dass, wenn technische Ineffizienz in großem Umfang in den Daten vorhanden ist, die ML Schätzung bessere Ergebnisse zeigt.⁵¹

⁵¹ Vgl. Coelli et al. (2003: 187).

Die ML-Methode ist ein sog. Punktschätzverfahren, das anders als die meisten anderen Techniken nicht vom arithmetischen Mittel sondern vom häufigsten Wert (Modus) einer Verteilung ausgeht. Grundgedanke des Verfahrens ist die Bestimmung der gesuchten Regressionsparameter ausgehend von einer Zufallsvariablen x , deren Dichte- bzw. Wahrscheinlichkeitsfunktion $f(x)$ von einem Parameter q abhängt. Liegt eine einfache Zufallsstichprobe mit n Realisationen vor, so lässt sich eine solche Dichtefunktion formulieren; maximiert man diese Funktion in Abhängigkeit von q (bildet die erste Ableitung nach q und setzt diese gleich Null), so resultiert eine Maximum-Likelihood-Schätzung. Es wird also der Wert von q gesucht, bei dem die Stichprobenwerte die größte Wahrscheinlichkeit haben. Unter der Annahme, dass der Effizienzterm und die Beobachtungen x für einzelne Firmen unabhängig voneinander sind, lässt sich zeigen, dass sog. „maximum-likelihood“-Schätzer (ML) effiziente Ergebnisse auch für β_0 erzeugen. Jedoch sind diese abhängig von der getroffenen Verteilungsannahme.⁵² Im Prinzip wird hierbei die Schätzgleichung wie eine zu maximierende Funktion (ML-Funktion) behandelt und die Formeln, nach denen die Schätzparameter berechnet werden, ergeben sich aus den Maximierungsbedingungen erster Ordnung. Für jede angenommene Verteilung sind daher neue Optimalbedingungen zur Maximierung der ML-Funktion zu bestimmen.

Da diese Ausdrücke sehr umfangreich werden, soll auf die komplette Darstellung hier verzichtet werden. Die wirkenden Mechanismen werden jedoch deutlich, wenn man den Ausdruck für die Dichtefunktion einer gestutzt-normalverteilten Variable u in Gleichung 4-10 analysiert:

$$(4-10) \quad f_{u_i}(u) = \frac{\exp\left[-\frac{1}{2}\frac{(u-\mu)^2}{\sigma_u^2}\right]}{(2\pi)^{\frac{1}{2}}\sigma_u^2\left[1-\Phi\left(-\frac{\mu}{\sigma_u}\right)\right]} \quad \forall \mu \geq 0$$

Gleichung (4-10) umfasst zunächst einmal die Halbnormal-Verteilung als Sonderfall für $\mu=0$, wobei Φ für die Dichtefunktion einer standard-normalverteilten Zufallsvariable steht. Im Normalfall wird jedoch davon ausgegangen, dass der Mittelwert der u_i einen positiven Wert annimmt. Dann werden aber auch die Schätzparameter einer Funktion wie (4-9) von diesen Verteilungsüberlegungen getrieben, denn der Mittelwert und die Standardabweichung der Effizienzstörterme wirken dann nicht nur auf ihre eigene Dichtefunktion; vielmehr beeinflussen sie insbesondere die Lage von β_0 und der anderen Regressionskoeffizienten.

Für populäre Verteilungen, wie die Halbnormal-Verteilung oder die gestutzte Normalverteilung existieren auf dem Markt für Analysesoftware wiederum maßgeschneiderte Lösungen, die bezüglich der genannten Verteilungen die

⁵² Für den Fall normalverteilter Effizienzterme gibt es keinen messtechnisch bedeutsamen Unterschied zwischen OLS-Schätzern und ML-Schätzern. Ist aber für einen Teil der gegebenen Beobachtungen tatsächlich $u_i > 0$ erfüllt, so verbessern sich die Werte der Likelihood-Funktion unter dieser Annahme gegenüber $u_i = 0$.

Effizienzwerte für einzelne Unternehmen in der Schätzung ermitteln.⁵³ Dabei wird häufig eine Reparametrisierung verwandt, die Battese und Corra entwickelt haben:⁵⁴

$$(4-11) \sigma_s^2 \equiv \sigma_u^2 + \sigma_v^2 \text{ und } \gamma \equiv \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} = \frac{\sigma_u^2}{\sigma_s^2}$$

Der Vorteil von γ in Gleichung (4-11) besteht darin, dass es nur Werte zwischen 0 und 1 annehmen kann und damit letztlich angibt, wie weit sich die SFA-Schätzung von dem Ergebnis einer OLS-Analyse entfernt. Geht γ gegen 1, so dominiert die einseitige technische Ineffizienz die symmetrischen Residuen v , sodass Gleichung (4-9) letztlich eine deterministische Produktionsfunktion ist. Umgekehrt gilt, dass alle Abweichungen rein zufällig und nicht systematischer Natur sind.

Dabei wird in ähnlicher Weise vorgegangen wie im MOLS-Verfahren, da zunächst eine Kleinstquadrat-Schätzung durchgeführt wird, deren Schätzparameter bis auf β_0 und σ_s^2 unverzerrt sind. In einem zweiten Schritt wird die Likelihoodfunktion für eine Anzahl von Werten von γ zwischen Null und eins ausgewertet, wobei β_0 und σ_s^2 korrigiert werden gemäß:

$$(4-12) \sigma_s^2 = \sigma_{OLS}^2 [\pi(T-K)] / [T(\pi - 2\hat{\gamma})] \text{ und}$$

$$(4-13) \hat{\beta}_0 = \hat{\beta}_{0(OLS)} + \sqrt{\frac{2\hat{\gamma}\hat{\sigma}_s^2}{\pi}}.$$

Für die übrigen β_i werden weiterhin die Werte aus der ursprünglichen OLS Schätzung verwendet. Beginnend mit den besten Schätzwerten aus Schritt zwei werden in einem iterativen Prozess nach Davidon-Fletcher-Powell abschließend die ML Schätzwerte als globale Maxima der Likelihoodfunktion bestimmt.

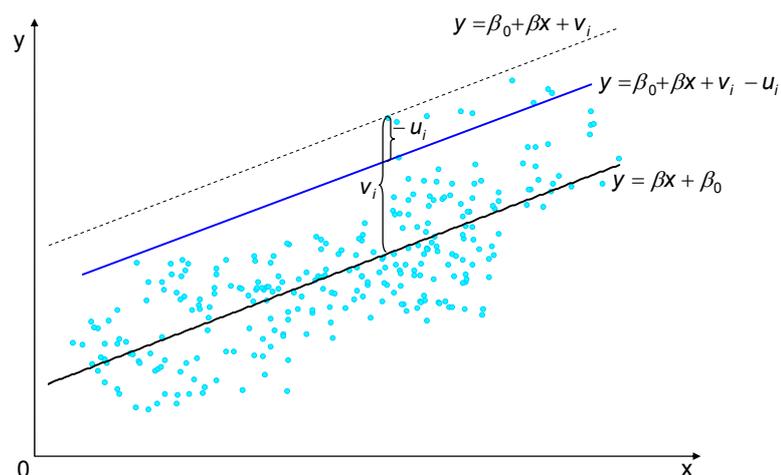
Es ergibt sich ein ähnliches Bild, wie wir es bereits aus Abbildung 4-3 kennen. Die untere Gerade in Abbildung 4-4 ist das Ergebnis einer normalen OLS Schätzung. Während COLS und MOLS weiterhin die Normalverteilungsannahme aufrecht erhalten, erweitert die SFA die Betrachtung insofern, als dass sie die Effizienzvariablen nicht mehr als normalverteilt annimmt, sondern vielmehr davon ausgeht, dass es signifikante Produktions- bzw. Kostenabweichungen in eine bestimmte Richtung gibt (z.B. eine gestutzte „Normal“-Verteilung). In diesem Fall können OLS-Schätzer nicht mehr für alle Parameter der Regression effizient sein. Betroffen hiervon ist vor allem das Absolutglied der Regressionsgerade β_0 , da die relative Effizienz der die Grenze bildenden Unternehmung die Lage der Gerade einseitig beeinflusst – sie liegt daher,

⁵³ Hierzu zählen das ebenfalls von Coelli entwickelte Frontier 4.1 (erhältlich unter: <http://www.uq.edu.au/economics/cepa/frontier.htm>) oder die Statistiksoftware LIMDEP, vgl. hierzu auch Sena (1999).

⁵⁴ Vgl. Battese und Corra (1977: 171).

wie in Abbildung 4-4 zu sehen ist, nicht länger in der Mitte der Punktwolke (mittlere Gerade). Grob gesprochen beinhaltet die SFA die beiden unter COLS und MOLS beschriebenen Mechanismen. Zum einen wird – analog zu COLS - unter der alleinigen Verwendung der Störterme v_i eine effiziente Grenze bestimmt (obere Gerade in Abbildung 4-4). Zum anderen wird – analog zu MOLS - unter Zuhilfenahme der als technische Ineffizienz zu interpretierenden u_i die relevante Regressionsgrade durch die Punktwolke der Beobachtungen gelegt. Im Endeffekt vereinnahmt die Stochastic Frontier Analysis die Vorteile der beiden vorhergehenden Methoden ohne ihre jeweiligen Schwächen und ist daher gegenüber COLS und MOLS zu präferieren. Gegenüber der DEA hat die SFA den Vorteil, dass Datenungenauigkeiten über die Störterme aufgefangen werden können. Während die DEA sehr sensibel auf Ausreißer reagiert, werden die Ergebnisse bei SFA aber oft durch die zugrunde liegenden Annahmen über die Störterme beeinflusst. So ergeben Exponentialverteilungen tendenziell robustere Ergebnisse als die Halbnormalverteilung.

Abbildung 4.4: SFA Produktionsschätzung



4.3.1.2 Produktions-, Kosten- und Profitschätzungen

Ähnlich wie die DEA erlaubt auch die SFA dem Anwender letztlich aus einer Vielzahl möglicher Ansätze zu wählen. Zu erwähnen ist hier vor allem die von der Dualitätsannahme getriebene direkte Schätzung von Kostenfunktionen bei ausreichend definierten Produktionsannahmen. So bedingen konstante Skalenerträge über den gesamten Bereich der relevanten Nachfrage konstante Grenz- und Durchschnittskosten, wogegen fallende Skalenerträge zu steigenden Grenz- und Durchschnittskosten führen würden.

Analog zur DEA ließe sich die SFA ebenfalls outputorientiert durchführen unter Verwendung von Schätzungen der Gewinnfunktionen, um Aussagen über die allokativen Effizienz zu erlauben. Fraglich erscheint erneut, inwiefern dieser Ansatz bei regulierten Unternehmen wie den Netzbetreibern gerechtfertigt ist. Grundsätzlich dürfte davon ausgegangen werden können, dass die Netzbetreiber zumindest in Teilen in der Position eines Monopolisten sind und diese daraus resultierende Macht für preissetzendes Verhalten auch nutzen. Eine Anwendung der SFA unter diesen Bedingungen würde zu Ergebnissen führen, die nicht mehr ökonomisch aussagekräftig wären. Zwar kann die Methode auch für diesen Fall theoretisch angepasst werden, allerdings verliert dann die Dualitätsprämisse ihre Gültigkeit.⁵⁵ Ferner bestünde nun die Notwendigkeit, Nachfragefunktionen neben Informationen über Technologien explizit aufzunehmen, was den Regulierer in der Praxis häufig vor enorme Probleme aufgrund der fehlenden Datenbasis stellen dürfte. Daher ist es unserer Ansicht nach ratsam, sich auf die Inputseite zu konzentrieren, zumal das Management der Netzbetreiber – zumindest kurzfristig – mehr Kontrolle über diese Seite des Produktionsprozesses haben dürfte.

⁵⁵ Vgl. Kumbhakar und Lovell (2003: 212ff.).

5 Zusammenspiel von DEA und SFA

Bei über 900 Netzgesellschaften in Deutschland erscheint die Technik des Benchmarkings insgesamt geeignet, verlässliche Aussagen über relative Effizienzstandards zwischen den Unternehmen zu produzieren. Die Anzahl dürfte groß genug sein, um hinreichend aussagekräftige Vergleichsergebnisse zu erhalten. Nach der vorangehenden Beschreibung der verschiedenen Möglichkeiten zum Benchmarking von Netzgesellschaften stellt sich die Frage, auf welche Methodik sich ein Regulierer konzentrieren sollte. Im Laufe des Textes sind bereits einige Hinweise gegeben worden, die hier kurz zusammenfassend dargestellt werden sollen (siehe auch Tabelle 5.1). Ferner soll noch ein Ausblick auf eine – unserer Meinung nach – aussichtsreiche Komplementarität zweier Methoden, der DEA und der SFA, gegeben werden.

Die traditionellen Ansätze zur Schätzung von Effizienzgrenzen (OLS, COLS und MOLS) können dabei allerdings als nicht problemadäquat eingestuft werden, da sie – wie in Abschnitt 4.3.1.1 ausgeführt – durch die aus der gleichen Methodenklasse der stochastischen Ansätze stammende SFA dominiert werden. Hinsichtlich der drei anderen Verfahren (PIN, SFA und DEA) ist jedoch zu sagen, dass es nicht die „richtige“ Methode gibt. Bei der Auswahl des Ansatzes spielen vielmehr folgende Faktoren eine entscheidende Rolle:⁵⁶

- Datenverfügbarkeit,
- Bedeutung der Datenstreuung und
- Ziel der Analyse.

Allgemein ist auf die Konsistenz der verwendeten Daten zu achten, insbesondere bei Zeitreihen. Soll die Veränderung der TFP über die Zeit bestimmt werden, dürfte es in einigen Fällen schwierig sein, Paneldaten auf Unternehmensebene zu erhalten. In diesem Fall bietet sich die Anwendung der PIN an, bei der auf aggregierte Daten auf Sektorebene zurückgegriffen werden kann, die häufig aus den offiziellen Statistiken entnommen werden können. Allerdings ist auch hier zu beachten, ob es Brüche in den erhobenen Daten gibt, z.B. aufgrund von Änderungen in den Messungen bzw. Definitionen der zugrunde liegenden Parameter. Bei SFA und DEA können diese Probleme zumindest teilweise umgangen werden, wenn auf rein physische Daten zurückgegriffen wird, die in der Regel einer geringeren Gefahr der Inkonsistenz unterliegen dürften als Preisangaben (Indexierung).

Ferner stellt sich die Frage, wie verlässlich die verfügbaren Daten sind. Während unter DEA angenommen wird, dass es keine Datenungenauigkeiten gibt, werden diese im Rahmen von SFA explizit berücksichtigt. Liegen Daten minderer Güte vor, so sollte daher eher auf SFA zurückgegriffen werden. Es sei nicht verschwiegen, dass auch

⁵⁶ Vgl. Coelli et al. (2003: 99ff.).

diese Methode anfällig gegenüber Datenfehlern ist, da zur Dekomposition der Residuen in Störterm und Ineffizienz ein Schiefemaß verwendet wird, das wiederum dem Einfluss von Ausreißern unterliegt. Ein weiterer Vorteil von SFA ist, dass die Verlässlichkeit bzw. Robustheit der Effizienzergebnisse mittels Konfidenzintervallen einer Überprüfung unterzogen werden kann. Diesem Mangel bei der DEA ist allerdings in der Zwischenzeit begegnet worden (z.B. über Bootstrapping, siehe auch Abschnitt 3.5).

Wird beabsichtigt, unternehmensspezifische X-Faktoren im Rahmen einer Price Cap Regulierung zu bestimmen, ist PIN aufgrund des höheren Aggregationsgrades der Daten ungeeignet. Allerdings erscheint auch eine rein mechanische Übertragung der Ergebnisse aus SFA oder DEA nicht ratsam. Einem Unternehmen, das größere Investitionsaktivitäten plant (z.B. allein aufgrund der Altersstruktur der Anlagen), wird es leichter fallen als anderen Betreibern, die Vorgaben zu erfüllen.

Grundsätzlich sollten DEA und SFA nicht als zueinander konfliktär betrachtet werden. Im Gegenteil bieten die zwei verfügbaren Methoden eher eine Möglichkeit an, die ermittelten jeweiligen Effizienzergebnisse auf ihre Plausibilität zu untersuchen (Korrelation der Ergebnisse). Fallen also DEA-Effizienz und SFA-Werte deutlich auseinander, so sind beide Untersuchungen nochmals auf Fehler abzuklopfen. Einerseits könnten hier Datenfehler zugrunde liegen – wahrscheinlicher aber ist, dass entweder die angenommene Produktions- oder Kostenfunktion zu restriktiv ist (SFA) oder die Bestimmung von relevanten Inputs und Outputs ganz allgemein überdacht werden sollte (DEA).

Möglich könnte es aber auch sein, die Ergebnisse der einen Analyse im Rahmen der anderen zu verwenden. Zu denken ist hier insbesondere daran, die Ergebnisse der DEA in einer ökonometrischen Analyse weiter zu verwenden. So ließe sich zur Schätzung der Produktionsfunktion z.B. ein Dummy heranziehen, der allen DEA-effizienten Unternehmen den Wert 1 und allen anderen den Wert 0 zuordnet. In diesem Fall lässt sich die Signifikanz der Resultate der linearen Optimierung dann ökonometrisch hinterfragen.

Abschließend sei noch angemerkt, dass es aufgrund der Monopolsituation bei den Netzbetreibern anzuraten ist, sowohl bei DEA als auch bei SFA auf inputorientierte Ansätze zurückzugreifen, die zwar keine Aussagen über die allokativen Effizienz auf der Outputseite zulassen, welche allerdings bei preissetzendem Verhalten auch von geringer Aussagekraft ist. Ferner kann dem Problem der Strukturmerkmale, auf die die Unternehmen keinen Einfluss haben, bei der DEA über die Verwendung von Umweltvariablen begegnet werden, wobei vor einer inflationären Handhabung gewarnt wird, da andernfalls die Ergebnisse wiederum ihre Aussagekraft verlieren.

An diesem Punkt sollte in der Zukunft angesetzt werden, indem Probedaten für eine Reihe von Input- und Outputvariablen der Netzgesellschaften erhoben werden, um die Bedeutung und Wirkungsweise der verschiedenen Kostentreiber zu evaluieren und konkrete Aussagen über deren zukünftige Verwendung im Rahmen eines regulatorischen Benchmarking Verfahrens treffen zu können.

Tabelle 5.1 : Die drei Benchmarking-Verfahren im Vergleich

	Price-based index numbers (PIN)	Data envelopment analysis (DEA)	Stochastic frontier analysis (SFA)
Beschreibung	Annäherung traditioneller Indexpzahlen an die TFP Messung. Preise werden als Gewichte benutzt. Gewöhnlich werden die Törnqvist oder Fisher Formel benutzt.	Lineare Optimierungsmethode, die eine nicht-parametrische Produktionsgrenze konstruiert, indem stückweise eine lineare Oberfläche auf die Datenpunkte gesetzt wird.	Ökonometrische Methode welche eine Produktionsfunktion der Form $y=f(x)+v-u$ schätzt, wobei y der Output ist, $f(x)$ sind alle Inputs, v ist ein Störterm für zufällige Abweichungen, und beschreibt die technische Ineffizienz. Eine Kostengrenze (kurzfristig oder langfristig) oder eine Abstandsfunktion kann stattdessen benutzt werden.
Benötigte Daten	Mengen- und Preisdaten von Inputs und Outputs zweier oder mehrerer Firmen, idealerweise für mehrere Jahre.	Mengendaten von Input und Output für eine Auswahl von Firmen, idealerweise für mehrere Jahre. Falls jedoch Preisdaten erhältlich sind, können diese ebenfalls zur Berechnung der allokativen Effizienz genutzt werden.	<ul style="list-style-type: none"> • <i>Produktionsgrenze / Abstandsfunktion:</i> Mengendaten von Inputs und Outputs einer Auswahl von Firmen, idealerweise über mehrere Jahre. • <i>langfristige Kostengrenze:</i> Gesamtkosten, Input-Preise, Output-Mengen • <i>kurzfristige Kostengrenze:</i> variable Kosten, variable Inputpreise, fixe Inputmengen und Outputmengen
Vorteile	<ul style="list-style-type: none"> • Studie mit nur zwei Beobachtungen möglich. • Reproduzierbar und transparent. • Berechnet steuerbare Effizienz. 	<ul style="list-style-type: none"> • Identifiziert eine Gruppe gleichrangiger Firmen (effiziente Firmen mit ähnlichem Input und Output Kombinationen) für jede ineffiziente Firma. • Einfache Bewältigung mehrerer Outputs. • Setzt keine funktionale Form für die Grenze oder eine aufgeteilte Form für den ineffizienten Störterm voraus. 	<ul style="list-style-type: none"> • Versuche der Betrachtung der Störanfälligkeit • Umweltbedingte Variablen sind leichter abzuhandeln • Ausreißer sind leichter zu identifizieren • Kostengrenze und Abstandsfunktionen können mit mehreren Outputs arbeiten
Nachteile	<ul style="list-style-type: none"> • Preisinformationen werden benötigt. • Keine Zerlegung der TFP Messung in Bestandteile möglich. 	<ul style="list-style-type: none"> • Kann von Störfaktoren beeinflusst werden. • Traditioneller Hypothesentest ist nicht möglich. • Benötigt eine große Auswahl für eine stabile Schätzung, welche in den ersten Jahren u.U. nicht leicht zu beschaffen sind. 	<ul style="list-style-type: none"> • Zerlegung der Störterme in Stör- und Effizienzbestandteile kann beeinflusst werden von speziellen Streuformen und von der zugehörigen Annahme, dass Fehlerschiefe ein Indikator von Ineffizienz ist. • Benötigt eine große Auswahl für eine stabile Schätzung, welche in den ersten Jahren u.U. nicht leicht zu beschaffen sind.

Annex 1: Preisindex-Formeln (PIN)

Üblicherweise werden in regulatorischen Fragestellungen, die sich mit TFP befassen, Preisinformationen als Gewichte verwendet, so dass

$$(A1-1) \Delta TFP_1/TFP_0 = \frac{\sum_{j=1}^J a_{j1} \cdot Y_{j1}}{\sum_{k=1}^K b_{k1} \cdot X_{k1}} \bigg/ \frac{\sum_{j=1}^J a_{j1} \cdot Y_{j0}}{\sum_{k=1}^K b_{k1} \cdot X_{k0}} = \frac{\sum_{j=1}^J a_{j1} \cdot Y_{j1}}{\sum_{j=1}^J a_{j1} \cdot Y_{j0}} \bigg/ \frac{\sum_{k=1}^K b_{k1} \cdot X_{k1}}{\sum_{k=1}^K b_{k1} \cdot X_{k0}}$$

mit

a_{ji} : Preis für Output j in Periode i , $i=0,1$;

Y_{ji} : Menge von Produkt j in Periode i , $i=0,1$;

b_{ki} : Preis für Input k in Periode i , $i=0,1$;

X_{ki} : Menge von Produktionsfaktor k in Periode i , $i=0,1$.

einen Index für die totale Faktorproduktivität unter der Annahme beschreibt, die Preise seien zwischen Periode 0 und 1 veränderlich, wobei in Gleichung A1-1 die Preise der Periode 1 als Gewichte gewählt wurden. Der rechte Ausdruck ist dann der Quotient eines Output-Mengenindex mit einem Input Index, wobei beide Indices hier nach Paasche ermittelt werden. Würde man dagegen die Preise der Ausgangsperiode wählen, würde es sich um Indices nach Laspeyres handeln.⁵⁷ Da die Indices jeweils zwei Extrempunkte der tatsächlichen Veränderung messen (Basisperiode vs. aktuelle Periode), wird in der regulatorischen Praxis jedoch häufig der sog. Fisher Index verwandt, der sich aus dem geometrischen Mittel der beiden traditionellen Maßzahlen errechnet, wobei der Quotient zweier Fisher-Indices wiederum einen TFP-Index bildet.:

$$(A1-2) \text{ Fisher Output - Index} = \sqrt{\frac{\sum_{j=1}^J a_{j0} \cdot Y_{j1}}{\sum_{j=1}^J a_{j1} \cdot Y_{j1}} \cdot \frac{\sum_{j=1}^J a_{j1} \cdot Y_{j1}}{\sum_{j=1}^J a_{j0} \cdot Y_{j0}}}$$

⁵⁷ Aus der Theorie der Messung der Konsumentenpreise, bei der die Rolle der Gewichte von den Mengen eingenommen wird, ist bekannt, dass Paasche- und Laspeyres Indices im Zeitablauf voneinander abweichen und Messfehler produzieren. So überschätzen Laspeyres Indices – wie z.B. der HVPI – die Preisentwicklung, da sie das Substitutionsverhalten der Konsumenten nicht abbilden. Analog hierzu kann nicht ausgeschlossen werden, dass Mengenindices die Entwicklung der Inputeinsätze unvollständig abbilden, wenn infolge steigender Preise Inputs substituiert werden. Wegen der als konstant angenommenen Preise ist der Index dann im Vergleich zur realen Entwicklung verzerrt.

$$(A1-3) \text{ Fisher TFP - Index} = \frac{\sqrt{\frac{\sum_{j=1}^J a_{j0} \cdot Y_{j0}}{\sum_{j=1}^J a_{j1} \cdot Y_{j1}} \cdot \frac{\sum_{j=1}^J a_{j1} \cdot Y_{j1}}{\sum_{j=1}^J a_{j0} \cdot Y_{j0}}}}{\sqrt{\frac{\sum_{k=1}^K b_{k0} \cdot X_{k0}}{\sum_{k=1}^K b_{k1} \cdot X_{k1}} \cdot \frac{\sum_{k=1}^K b_{k1} \cdot X_{k1}}{\sum_{k=1}^K b_{k0} \cdot X_{k0}}}}$$

Der Fisher Index hat eine Reihe theoretischer Eigenschaften, die für seine Verwendung sprechen⁵⁸, jedoch zieht die Literatur in jüngerer Zeit einen anderen Index vor, dessen produktionstheoretischer Hintergrund noch flexibler interpretiert werden kann.

$$(A1-4) \text{ Törnquist Outputindex} = \prod_{i=1}^L \left[\frac{Y_{i1}}{Y_{i0}} \right]^{\frac{\omega_{i1} + \omega_{i0}}{2}} \quad \text{mit } \omega_{it} = \frac{p_{it} \cdot X_{it}}{\sum_{i=1}^L p_{it} \cdot X_{it}}$$

Dem Törnquist-Index liegt eine sog. Translog-Technologie zugrunde; und auch der zugehörige TFP-Index selbst wird häufig in logarithmierter Form berechnet.⁵⁹

$$(A1-5) \ln(TFP_1/TFP_0) = 0,5 \sum_{l=1}^L [(\omega_{l1} + \omega_{l0}) \cdot (y_{l1} - y_{l0})] - 0,5 \sum_{m=1}^M [(\vartheta_{m1} + \vartheta_{m0}) \cdot (x_{m1} - x_{m0})]$$

⁵⁸ So impliziert seine Verwendung in der TFP-Analyse die Existenz einer **quadratischen** Technologie (etwa Cobb Douglas), wogegen die traditionellen Indices letztlich lineare Technologien abbilden. Zudem erfüllt der Index eine Reihe sog. Test-Kriterien, die ebenfalls von Fisher entwickelt wurden. Hierzu gehören die Proportionalität, die Invarianz gegenüber Maßeinheiten, die Reversibilität in der Zeit und andere.

⁵⁹ Vgl. Christensen et al. (1971). Siehe hierzu auch Abschnitt 4.1 Alternative Spezifikationen der Produktionsfunktion. Grundsätzlich stellt eine Translog die logarithmierte Form eines Taylor-Polynoms zweiten Grades dar, was im Prinzip bedeutet, dass sie z.B. eine Cobb-Douglas Produktionsfunktion abbilden könnte, aber nicht muss.

Annex 2: Technik der linearen Programmierung bei DEA

Der DEA liegt eine Verallgemeinerung der eindimensionalen Produktivitätsanalyse zugrunde (vgl. Kapitel 2). Will man Produktivitätskennzahlen verallgemeinern, sodass mehrere Inputs und Outputs berücksichtigt werden können, ergibt sich das bereits aus der Diskussion der PIN bekannte Problem der Gewichtungen. Während die Gewichte bei Preis- oder Mengenindices exogen bestimmt werden, ermittelt die DEA die Gewichte endogen und zwar so, dass der Quotient aus gewichteten Out- und Inputs maximiert wird. Dabei ist u ein $M \times 1$ -Vektor der Outputgewichte, während v einen $K \times 1$ -Vektor der Inputgewichte bildet.

$$(A2-1) \max h = \frac{u' y_i}{v' x_i} \quad \text{s.t.} \quad \frac{u' y_j}{v' x_j} \leq 1 \quad \forall j = 1, 2, \dots, N; \quad u, v \geq 0$$

Die Nebenbedingungen in Gleichung (A2-1) sorgen dafür, dass der Effizienzwert für keine Unternehmung größer als 1 wird, während gleichzeitig sichergestellt ist, dass allen Beobachtungen ein möglichst großer Wert zugewiesen wird – die Unternehmen werden gewissermaßen in ein möglichst vorteilhaftes Licht gerückt. Allerdings existieren für Gleichung (A2-1) – ein nichtlineares Problem – unendlich viele Lösungen, da für bestimmte optimale u und v auch immer deren Vielfache als Lösung akzeptiert werden, was die praktische Anwendbarkeit einschränkt. Diese Problematik lässt sich jedoch umgehen, indem eine zusätzliche Restriktion eingeführt wird, die den Nenner der Zielfunktion eliminiert ($v' x_i = 1$). Gleichzeitig werden die n Nebenbedingungen mit dieser neuen Restriktion multipliziert und nach Null aufgelöst, wobei die geänderte Notation darauf hinweist, dass es sich um ein neues lineares Programmierungsproblem handelt.

$$(A2-2) \max (\mu' y_i) \quad \text{s.t.} \quad v' x_i = 1; \quad \mu' y_j - v' x_j \leq 0 \quad \forall j = 1, 2, \dots, N; \quad \mu, v \geq 0$$

Gleichung (A2-2) muss für jede Beobachtung/Unternehmung – also insgesamt n mal – gelöst werden, dabei bilden die Multiplikatoren (μ, v) die Variablen des Maximierungsproblems, wobei für jede Unternehmung verlangt wird, dass die Summe ihrer gewichteten Outputs kleiner sei als die gewichtete Summe der Inputs.

Eine spezielle Eigenschaft bestimmter linearer Optimierungsprobleme – ihre **Dualität** – erlaubt es nun, Gleichung (A2-2) nochmals zu verändern, umso den Rechenaufwand weiter zu reduzieren. Die Dualität beruht dabei auf folgender Überlegung: Zum Lösen des LP werden Zeilen mit bestimmten Konstanten multipliziert und sodann zur Zielfunktion addiert und zwar solange, bis das Optimaltableau erreicht ist (Simplexmethode). Anstatt diesen Prozess Schritt für Schritt durchzuführen, lässt sich aber auch gleich nach denjenigen positiven Konstanten λ_j fragen, mit denen die Zeilen multipliziert und zur Zielzeile addiert werden müssten, um das Problem in einem Schritt zu lösen. Wählt man diese Form des Herangehens, so dreht sich das LP um, d.h. aus einer Maximierung wird eine Minimierung, die Koeffizientenmatrix (hier Inputs und

Outputs) wird transponiert und die Elemente der Zielfunktion werden zu Variablen der Nebenbedingungen.

$$(A2-3) \min \theta \quad s.t. \quad -y_i + Y\lambda \geq 0; \quad \theta x_i - X\lambda \geq 0; \quad \lambda \geq 0$$

Gleichung (A2-3) lässt sich aus ökonomischer Sicht so interpretieren, dass der Output y_i einer bestimmten Unternehmung konstant gehalten wird, während ein Wert θ gesucht wird, der angibt auf welchen Bruchteil von 100 sich die Inputs x_i auf einem Strahl reduzieren ließen, ohne dass es zu einer Reduktion der Produktionsmenge käme. Dabei bilden Inputwerte anderer Unternehmungen X eine Hülle, deren Lage bestimmt, in welchem Umfang es möglich ist, den Einsatz der Faktoren zu reduzieren. θ bezeichnet insofern den Effizienzwert, der einer bestimmten Unternehmung zugewiesen wird, wenn wie hier eine Input-orientierte DEA erfolgt. Gleichung (A2-3) hat zudem den Vorteil, dass sie nicht n Restriktionen, sondern nur $K+M$ Nebenbedingungen hat. Ein derartiges lineares Optimierungsproblem lässt sich mit einer Vielzahl von Softwarepaketen (Excel, MatLab, Shazam) lösen. Zusätzlich gibt es eine ganze Reihe spezieller DEA-Programme, die diese und alle weiteren Kalkulationen, die bei einer Reduktion der Abstraktion auftreten, lösen können.⁶⁰

⁶⁰ Hierzu gehört z.B. das Programm von Cooper, Seiford und Tone (2000) oder das Programm DEAP (2.1) von Coelli (1996). Letztere Software ist unter <http://www.uq.edu.au/economics/cepa/deap.htm> frei im Internet verfügbar.

Literaturverzeichnis

- Adler, N., L. Friedman und Z. Sinuany-Stern (2002), Review of ranking methods in the data envelopment analysis context, *European Journal of Operational Research* 140, 249-265.
- Ajodhia, V., K. Petrov und G.C. Scarsi (2003), Benchmarking and its Applications, *Zeitschrift für Energiewirtschaft* 27, 261-274.
- Battese, G. E. und G.S. Corra (1977), Estimation of a Production Frontier Model: With Application to the Pastoral Zone of Eastern Australia, *Australian Journal of Agricultural Economics* 21 (3), 169-179.
- Christensen, L.R., D.W. Jorgensen und L.J. Lau. (1971), Conjugate Duality and the Transcendental Logarithmic Production Function, *Econometrica* 39, 255-256.
- Coelli, T.J. (1995), A multi-stage methodology for the solution of oriented DEA models, *Operations Research Letters* 23, 143-149.
- Coelli, T.J. (1996), A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program, CEPA Working Paper 96/08, Department of Econometrics, University of New England, Armidale.
- Coelli, T. J., D.S. Prasada Rao und G.E. Battese (1998), *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Boston.
- Coelli, T., A. Estache, S. Perelman und L. Trujillo (2003), *A Primer on Efficiency Measurement for Utilities and Transport Regulators*, World Bank Institute Development Studies, Washington, D.C..
- Cooper, W., L. Seiford und K. Tone (2000), *Data Envelopment Analysis*, Boston.
- Färe, R., S. Grosskopf, M. Norris und Z. Zhang (1994), Productivity Growth, Technical Progress and Efficiency Change in Industrialized Countries, *American Economic Review* 84(1), 66-83.
- Ferrier, G.D. und C.A.K. Lovell (1990), Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence, *Journal of Econometrics* 46, 229-245.
- Hense, A. und D. Schäffner (2004), *Regulatorische Aufgaben im Energiebereich – ein europäischer Vergleich*, wik Diskussionsbeiträge Nr. 254, Bad Honnef.
- Jamasb, T. und M. Pollitt (2001), Benchmarking and Regulation: International Electricity Experiences, *Utilities Policy*, 9, 107-130.
- Jamasb, T. und M. Pollitt (2003), International Benchmarking and Yardstick Regulation: An Application to European Electricity Distribution Utilities, *Energy Policy* 31, 1609-1622.
- Kirjavainen, T. und H.A. Loikkanen (1998), Efficiency differences of Finnish senior secondary schools: an application of DEA and Tobit analysis, *Economics of Education Review* 17(4), 377-394.
- Kumbhakar, S.C. und C.A.K. Lovell (2003), *Stochastic Frontier Analysis*, Cambridge, U.K..
- Mas-Colell, A., M.D. Whinston und J.R. Green (2004), *Microeconomic Theory*, Oxford University Press, Oxford.

Olesen, OB (2004), Comparing and Combining Two Approaches for Chance Constrained DEA, Working paper submitted to the Journal of Productivity Analysis.

Sena, V. (1999), Stochastic Frontier Estimation: A Review of the Software Options, Journal of Applied Econometrics 14, 579-586.

Simar, L. und P. Wilson (2000a), A general methodology for bootstrapping in non-parametric frontier models, Journal of Applied Statistics 27, 779-802.

Simar, L. und P. Wilson (2000b), Statistical Interference in Nonparametric Frontier Models: The State of the Art, Journal of Productivity Analysis 13, 49-78.

Als "Diskussionsbeiträge" des Wissenschaftlichen Instituts für Kommunikationsdienste sind zuletzt erschienen:

- Nr. 180: Cara Schwarz-Schilling:
Nummernverwaltung bei Wettbewerb in der Telekommunikation, Dezember 1997
also available in English as
Numbering Administration in Telecommunications under Competitive Conditions
- Nr. 181: Cornelia Fries:
Nutzerkompetenz als Determinante der Diffusion multimedialer Dienste, Dezember 1997
- Nr. 182: Annette Hillebrand:
Sicherheit im Internet zwischen Selbstorganisation und Regulierung - Eine Analyse unter Berücksichtigung von Ergebnissen einer Online-Umfrage, Dezember 1997
- Nr. 183: Lorenz Nett:
Tarifpolitik bei Wettbewerb im Markt für Sprachtelefondienst, März 1998
- Nr. 184: Alwin Mahler:
Strukturwandel im Bankensektor - Der Einfluß neuer Telekommunikationsdienste, März 1998
- Nr. 185: Henrik Hermann:
Wettbewerbsstrategien alternativer Telekommunikationsunternehmen in Deutschland, Mai 1998
- Nr. 186: Ulrich Stumpf, Daniel Tewes:
Digitaler Rundfunk - vergleichende Betrachtung der Situation und Strategie in verschiedenen Ländern, Juli 1998
- Nr. 187: Lorenz Nett, Werner Neu:
Bestimmung der Kosten des Universaldienstes, August 1998
- Nr. 188: Annette Hillebrand, Franz Büllingen:
Durch Sicherungsinfrastruktur zur Vertrauenskultur: Kritische Erfolgsfaktoren und regulatorische Aspekte der digitalen Signatur, Oktober 1998
- Nr. 189: Cornelia Fries, Franz Büllingen:
Offener Zugang privater Nutzer zum Internet - Konzepte und regulatorische Implikationen unter Berücksichtigung ausländischer Erfahrungen, November 1998
- Nr. 190: Rudolf Pospischil:
Repositionierung von AT&T - Eine Analyse zur Entwicklung von 1983 bis 1998, Dezember 1998
- Nr. 191: Alfons Keuter:
Beschäftigungseffekte neuer TK-Infrastrukturen und -Dienste, Januar 1999
- Nr. 192: Wolfgang Elsenbast:
Produktivitätserfassung in der Price-Cap-Regulierung – Perspektiven für die Preisregulierung der Deutschen Post AG, März 1999
- Nr. 193: Werner Neu, Ulrich Stumpf, Alfons Keuter, Lorenz Nett, Cara Schwarz-Schilling:
Ergebnisse und Perspektiven der Telekommunikationsliberalisierung in ausgewählten Ländern, April 1999
- Nr. 194: Ludwig Gramlich:
Gesetzliche Exklusivlizenz, Universaldienstpflichten und "höherwertige" Dienstleistungen im PostG 1997, September 1999
- Nr. 195: Hasan Alkas:
Rabattstrategien marktbeherrschender Unternehmen im Telekommunikationsbereich, Oktober 1999
- Nr. 196: Martin Distelkamp:
Möglichkeiten des Wettbewerbs im Orts- und Anschlußbereich des Telekommunikationsnetzes, Oktober 1999
- Nr. 197: Ulrich Stumpf, Cara Schwarz-Schilling unter Mitarbeit von Wolfgang Kiesewetter:
Wettbewerb auf Telekommunikationsmärkten, November 1999
- Nr. 198: Peter Stamm, Franz Büllingen:
Das Internet als Treiber konvergenter Entwicklungen – Relevanz und Perspektiven für die strategische Positionierung der TIME-Player, Dezember 1999

- Nr. 199: Cara Schwarz-Schilling, Ulrich Stumpf:
Netzbetreiberportabilität im Mobilfunkmarkt – Auswirkungen auf Wettbewerb und Verbraucherinteressen, Dezember 1999
- Nr. 200: Monika Plum, Cara Schwarz-Schilling:
Marktabgrenzung im Telekommunikations- und Postsektor, Februar 2000
- Nr. 201: Peter Stamm:
Entwicklungsstand und Perspektiven von Powerline Communication, Februar 2000
- Nr. 202: Martin Distelkamp, Dieter Elixmann, Christian Lutz, Bernd Meyer, Ulrike Schimmel:
Beschäftigungswirkungen der Liberalisierung im Telekommunikationssektor in der Bundesrepublik Deutschland, März 2000
- Nr. 203: Martin Distelkamp:
Wettbewerbspotenziale der deutschen Kabel-TV-Infrastruktur, Mai 2000
- Nr. 204: Wolfgang Elsenbast, Hilke Smit:
Gesamtwirtschaftliche Auswirkungen der Marktöffnung auf dem deutschen Postmarkt, Mai 2000
- Nr. 205: Hilke Smit:
Die Anwendung der GATS-Prinzipien auf dem Postsektor und Auswirkungen auf die nationale Regulierung, Juni 2000
- Nr. 206: Gabriele Kulenkampff:
Der Markt für Internet Telefonie - Rahmenbedingungen, Unternehmensstrategien und Marktentwicklung, Juni 2000
- Nr. 207: Ulrike Schimmel:
Ergebnisse und Perspektiven der Telekommunikationsliberalisierung in Australien, August 2000
- Nr. 208: Franz Büllingen, Martin Wörter:
Entwicklungsperspektiven, Unternehmensstrategien und Anwendungsfelder im Mobile Commerce, November 2000
- Nr. 209: Wolfgang Kiesewetter:
Wettbewerb auf dem britischen Mobilfunkmarkt, November 2000
- Nr. 210: Hasan Alkas:
Entwicklungen und regulierungspolitische Auswirkungen der Fix-Mobil Integration, Dezember 2000
- Nr. 211: Annette Hillebrand:
Zwischen Rundfunk und Telekommunikation: Entwicklungsperspektiven und regulatorische Implikationen von Webcasting, Dezember 2000
- Nr. 212: Hilke Smit:
Regulierung und Wettbewerbsentwicklung auf dem neuseeländischen Postmarkt, Dezember 2000
- Nr. 213: Lorenz Nett:
Das Problem unvollständiger Information für eine effiziente Regulierung, Januar 2001
- Nr. 214: Sonia Strube:
Der digitale Rundfunk - Stand der Einführung und regulatorische Problemfelder bei der Rundfunkübertragung, Januar 2001
- Nr. 215: Astrid Höckels:
Alternative Formen des entbündelten Zugangs zur Teilnehmeranschlussleitung, Januar 2001
- Nr. 216: Dieter Elixmann, Gabriele Kulenkampff, Ulrike Schimmel, Rolf Schwab:
Internationaler Vergleich der TK-Märkte in ausgewählten Ländern - ein Liberalisierungs-, Wettbewerbs- und Wachstumsindex, Februar 2001
- Nr. 217: Ingo Vogelsang:
Die räumliche Preisdifferenzierung im Sprachtelefondienst - wettbewerbs- und regulierungspolitische Implikationen, Februar 2001
- Nr. 218: Annette Hillebrand, Franz Büllingen:
Internet-Governance - Politiken und Folgen der institutionellen Neuordnung der Domainverwaltung durch ICANN, April 2001
- Nr. 219: Hasan Alkas:
Preisbündelung auf Telekommunikationsmärkten aus regulierungsökonomischer Sicht, April 2001

- Nr. 220: Dieter Elixmann, Martin Wörter:
Strategien der Internationalisierung im Telekommunikationsmarkt, Mai 2001
- Nr. 221: Dieter Elixmann, Anette Metzler:
Marktstruktur und Wettbewerb auf dem Markt für Internet-Zugangsdienste, Juni 2001
- Nr. 222: Franz Büllingen, Peter Stamm:
Mobiles Internet - Konvergenz von Mobilfunk und Multimedia, Juni 2001
- Nr. 223: Lorenz Nett:
Marktorientierte Allokationsverfahren bei Nummern, Juli 2001
- Nr. 224: Dieter Elixmann:
Der Markt für Übertragungskapazität in Nordamerika und Europa, Juli 2001
- Nr. 225: Antonia Niederprüm:
Quersubventionierung und Wettbewerb im Postmarkt, Juli 2001
- Nr. 226: Ingo Vogelsang
unter Mitarbeit von Ralph-Georg Wöhrl
Ermittlung der Zusammenschaltungsentgelte auf Basis der in Anspruch genommenen Netzkapazität, August 2001
- Nr. 227: Dieter Elixmann, Ulrike Schimmel, Rolf Schwab:
Liberalisierung, Wettbewerb und Wachstum auf europäischen TK-Märkten, Oktober 2001
- Nr. 228: Astrid Höckels:
Internationaler Vergleich der Wettbewerbsentwicklung im Local Loop, Dezember 2001
- Nr. 229: Anette Metzler:
Preispolitik und Möglichkeiten der Umsatzgenerierung von Internet Service Providern, Dezember 2001
- Nr. 230: Karl-Heinz Neumann:
Volkswirtschaftliche Bedeutung von Resale, Januar 2002
- Nr. 231: Ingo Vogelsang:
Theorie und Praxis des Resale-Prinzips in der amerikanischen Telekommunikationsregulierung, Januar 2002
- Nr. 232: Ulrich Stumpf:
Prospects for Improving Competition in Mobile Roaming, März 2002
- Nr. 233: Wolfgang Kiesewetter:
Mobile Virtual Network Operators – Ökonomische Perspektiven und regulatorische Probleme, März 2002
- Nr. 234: Hasan Alkas:
Die Neue Investitionstheorie der Realoptionen und ihre Auswirkungen auf die Regulierung im Telekommunikationssektor, März 2002
- Nr. 235: Karl-Heinz Neumann:
Resale im deutschen Festnetz, Mai 2002
- Nr. 236: Wolfgang Kiesewetter, Lorenz Nett und Ulrich Stumpf:
Regulierung und Wettbewerb auf europäischen Mobilfunkmärkten, Juni 2002
- Nr. 237: Hilke Smit:
Auswirkungen des e-Commerce auf den Postmarkt, Juni 2002
- Nr. 238: Hilke Smit:
Reform des UPU-Endvergütungssystems in sich wandelnden Postmärkten, Juni 2002
- Nr. 239: Peter Stamm, Franz Büllingen:
Kabelfernsehen im Wettbewerb der Plattformen für Rundfunkübertragung - Eine Abschätzung der Substitutionspotenziale, November 2002
- Nr. 240: Dieter Elixmann, Cornelia Stappen unter Mitarbeit von Anette Metzler:
Regulierungs- und wettbewerbspolitische Aspekte von Billing- und Abrechnungsprozessen im Festnetz, Januar 2003
- Nr. 241: Lorenz Nett, Ulrich Stumpf unter Mitarbeit von Ulrich Ellinghaus, Joachim Scherer, Sonia Strube Martins, Ingo Vogelsang:
Eckpunkte zur Ausgestaltung eines möglichen Handels mit Frequenzen, Februar 2003
- Nr. 242: Christin-Isabel Gries:
Die Entwicklung der Nachfrage nach breitbandigem Internet-Zugang, April 2003

- Nr. 243: Wolfgang Briglauer:
Generisches Referenzmodell für die Analyse relevanter Kommunikationsmärkte – Wettbewerbsökonomische Grundfragen, Mai 2003
- Nr. 244: Peter Stamm, Martin Wörter:
Mobile Portale – Merkmale, Marktstruktur und Unternehmensstrategien, Juli 2003
- Nr. 245: Franz Büllingen, Annette Hillebrand:
Sicherstellung der Überwachbarkeit der Telekommunikation: Ein Vergleich der Regelungen in den G7-Staaten, Juli 2003
- Nr. 246: Franz Büllingen, Annette Hillebrand:
Gesundheitliche und ökologische Aspekte mobiler Telekommunikation – Wissenschaftlicher Diskurs, Regulierung und öffentliche Debatte, Juli 2003
- Nr. 247: Anette Metzler, Cornelia Stappen unter Mitarbeit von Dieter Elixmann:
Aktuelle Marktstruktur der Anbieter von TK-Diensten im Festnetz sowie Faktoren für den Erfolg von Geschäftsmodellen, September 2003
- Nr. 248: Dieter Elixmann, Ulrike Schimmel with contributions of Anette Metzler:
"Next Generation Networks" and Challenges for Future Regulatory Policy, November 2003
- Nr. 249: Martin O. Wengler, Ralf G. Schäfer:
Substitutionsbeziehungen zwischen Festnetz und Mobilfunk: Empirische Evidenz für Deutschland und ein Survey internationaler Studien, Dezember 2003
- Nr. 250: Ralf G. Schäfer:
Das Verhalten der Nachfrager im deutschen Telekommunikationsmarkt unter wettbewerblichen Aspekten, Dezember 2003
- Nr. 251: Dieter Elixmann, Anette Metzler, Ralf G. Schäfer:
Kapitalmarktinduzierte Veränderungen von Unternehmensstrategien und Marktstrukturen im TK-Markt, März 2004
- Nr. 252: Franz Büllingen, Christin-Isabel Gries, Peter Stamm:
Der Markt für Public Wireless LAN in Deutschland, Mai 2004
- Nr. 253: Dieter Elixmann, Annette Hillebrand, Ralf G. Schäfer, Martin O. Wengler:
Zusammenwachsen von Telefonie und Internet – Marktentwicklungen und Herausforderungen der Implementierung von ENUM, Juni 2004
- Nr. 254: Andreas Hense, Daniel Schöffner:
Regulatorische Aufgaben im Energiebereich – ein europäischer Vergleich, Juni 2004
- Nr. 255: Andreas Hense:
Qualitätsregulierung und wettbewerbspolitische Implikationen auf Postmärkten, September 2004
- Nr. 256: Peter Stamm:
Hybridnetze im Mobilfunk – technische Konzepte, Pilotprojekte und regulatorische Fragestellungen, Oktober 2004
- Nr. 257: Christin-Isabel Gries:
Entwicklung der DSL-Märkte im internationalen Vergleich, Oktober 2004
- Nr. 258: Franz Büllingen, Annette Hillebrand, Diana Rätz:
Alternative Streitbeilegung in der aktuellen EMVU-Debatte, November 2004
- Nr. 259: Daniel Schöffner:
Regulierungsökonomische Aspekte des informatorischen Unbundling im Energiebereich, Dezember 2004
- Nr. 260: Sonja Schölermann:
Das Produktangebot von Universaldienstleistern und deren Vergleichbarkeit, Dezember 2004
- Nr. 261: Franz Büllingen, Aurélie Gillet, Christin-Isabel Gries, Annette Hillebrand, Peter Stamm:
Stand und Perspektiven der Vorratsdatenspeicherung im internationalen Vergleich, Februar 2005
- Nr. 262: Oliver Franz, Marcus Stronzik:
Benchmarking-Ansätze zum Vergleich der Effizienz von Energieunternehmen, Februar 2005