

**BDEW Bundesverband
der Energie- und
Wasserwirtschaft e.V.**
Reinhardtstraße 32
10117 Berlin

GEODE
Magazinstraße 15-16
10179 Berlin

**VKU Verband kommunaler
Unternehmen e.V.**
Invalidenstraße 91
10115 Berlin

Stellungnahme

zur Vorstellung der Vergleichsparameter für den Effizienzvergleich der deutschen Elektrizitätsverteilernetzbetreiber für die 3. Regulierungsperiode am 25. Juli 2018

Berlin, 16. August 2018

Inhaltsverzeichnis

EINLEITUNG	3
ZUSAMMENFASSUNG DER KERNFORDERUNGEN	4
1. BEURTEILUNG DES GESAMTPROZESSES	6
2. DATENQUALITÄT UND -VALIDIERUNG	8
2.1 VERÖFFENTLICHUNG UND AUFBAU DES KTA-DATENSATZES	8
2.2 BEHÖRDLICHE PLAUSIBILISIERUNGSPRÜFUNGEN.....	9
2.3 ANALYSE DER DATENGRUNDLAGE FÜR DIE IN DER KONSULTATION VORGESTELLTEN PARAMETER.....	10
2.4 PLAUSIBILISIERUNG DES KTA-DATENSATZES DURCH LAGEPARAMETERANALYSE	11
2.5 TRANSPARENZ DER DATENGRUNDLAGEN	15
3 KOSTENTREIBERANALYSE UND EFFIZIENZMESSUNG	15
3.1 STARKER FOKUS AUF STATISTISCHE KOSTENTREIBERANALYSE	15
3.1.1 <i>Annahme der linearen Funktion</i>	16
3.1.2 <i>Bestimmung der „optimalen Größe“ eines Modells</i>	17
3.1.3 <i>Zu starke Gewichtung auf die Problematik der Multikollinearität</i>	20
3.1.4 <i>Problem der Pfadabhängigkeit und verzerrte Teststatistiken</i>	22
3.1.5 <i>Keine Berücksichtigung neuerer wissenschaftlich erprobter Ansätze</i>	24
3.2 WAHL DER VERTEILUNG DER INEFFIZIENZ BEI DER SFA	24
3.3 BERÜCKSICHTIGUNG UNTERSCHIEDLICHER VERSORGUNGS-AUFGABEN (HETEROGENITÄTEN)	25
3.4 UNTERSCHIEDE DEA UND SFA/PARAMETERWAHL.....	28
3.5 DURCHFÜHRUNG DER AUSREIßERANALYSE.....	30
3.6 SECOND-STAGE-ANALYSEN/SENSITIVITÄT	33
LITERATUR	35

Einleitung

Die Bundesnetzagentur (BNetzA) wird für alle Stromverteilernetzbetreiber im regulären Verfahren einen Effizienzvergleich zur Ermittlung der unternehmensindividuellen Effizienzwerte der dritten Regulierungsperiode durchführen und letztlich die unternehmensindividuellen Erlösobergrenzen festlegen.

Am 25. Juli 2018 fand auf Einladung der BNetzA ein Workshop mit der Branche zur Vorstellung der vorläufigen Parameter für den Effizienzvergleich der dritten Regulierungsperiode für Stromverteilernetzbetreiber statt.

Im Vorfeld der Veranstaltung wurde seitens der BNetzA am 28. Juni 2018 eine Blankotabelle mit dem Datensatz zur Kostentreiberanalyse (KTA) veröffentlicht. Die mit Daten befüllte Tabelle wurde am 24. Juli 2018 auf der BNetzA-Homepage¹ veröffentlicht.

Vertreter der BNetzA und des BNetzA-Beraterkonsortiums Swiss Economics, Sumicsid, dem Institut und Lehrstuhl für Elektrische Anlagen und Energiewirtschaft (IAEW/RWTH Aachen) und ein Vertreter von BjörnSEN Beratende Ingenieure präsentierten dabei die folgenden Unterlagen bzw. Inhalte:

- den Stand der Untersuchungen zu den Datenlieferungen und der Datenplausibilisierung;
- das bisherige Vorgehen bei der Kostentreiberanalyse mit Erläuterungen zur Parameterauswahl inkl. Vorstellung einer möglichen, allerdings nicht finalen Parameterliste;
- drei ausgearbeitete Benchmarking-Modelle auf Basis der zuvor dargestellten Untersuchungen;
- einen Vergleich der drei Benchmarking-Modelle unter Berücksichtigung verschiedener statistischer Gütekriterien.

Aus der von der BNetzA veröffentlichten KTA-Tabelle geht hervor, dass insgesamt 204 Stromverteilernetzbetreiber im regulären Verfahren am Effizienzbenchmarking der dritten Regulierungsperiode teilnehmen.

Die Branche hat die Möglichkeit, bis zum 16. August 2018 Stellung zu nehmen. Nach Berücksichtigung der Stellungnahmen im Rahmen der Konsultation ist die Durchführung einer Validierung und eines Effizienzvergleichs durch die BNetzA gemäß den vorgestellten Methoden (ggf. angepasst) geplant. Die Resultate sollen den Stromverteilernetzbetreibern im September 2018 zur Verfügung gestellt werden, der ausformulierte Anhörungsbericht im 4. Quartal 2018. Der Schlussbericht mit den finalen Resultaten wird unter Berücksichtigung der unternehmensindividuellen Stellungnahmen erarbeitet.

BDEW/VKU/GEODE nehmen zu den von der BNetzA veröffentlichten Unterlagen sowie zu den Erläuterungen der BNetzA und deren Beratern sowie zu den Diskussionen während der Veranstaltung nachfolgend Stellung.

¹ https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Netzentgelte/Strom/EffizienzvergleichVerteilernetzbetreiber/3RegP/3RegP_node.html;jsessionid=0A55D02F10F41A19B69D89865FE3208F

Zusammenfassung der Kernforderungen

Es ist sicherzustellen, dass die umfassende Versorgungsaufgabe der Stromverteilernetzbetreiber im Effizienzvergleich adäquat abgebildet wird. Nachfolgende Schritte sind für eine sachgerechte und ARegV-konforme Ausgestaltung des Effizienzvergleichs der 3. Regulierungsperiode für Stromverteilernetzbetreiber notwendig und wichtige Randbedingungen sind zu beachten:

Die Heterogenität der Aufgaben der Netzbetreiber ist bei der Ermittlung der Effizienzwerte zu berücksichtigen

Die ARegV schreibt in § 13 ARegV vor: „Durch die Auswahl der Vergleichsparameter sollen die strukturelle Vergleichbarkeit möglichst weitgehend gewährleistet sein und die Heterogenität der Aufgaben der Netzbetreiber möglichst weitgehend abgebildet werden.“ Diese gesetzliche Vorgabe der ARegV muss zwingend umgesetzt und in die Ermittlung der Effizienzwerte einbezogen werden. Hierbei gilt es zu beachten, dass Modelle mit einer höheren Anzahl an Vergleichsparametern die Heterogenität der Aufgaben von Netzbetreibern methodisch bedingt erheblich besser abbilden.

Kriterium der Multikollinearität nicht verwenden

Das Kriterium der Multikollinearität sollte nicht dafür verwendet werden, um einzelne Parameter auszuschließen, da das alleinige Abstellen auf die statistische Signifikanz bei der Auswahl von Vergleichsparametern nicht sachgerecht ist.

Ausreißeranalyse: Anpassung der Methoden notwendig

Es ist eine Ausreißeranalyse durchzuführen, die der ARegV gerecht wird. Insbesondere muss diese den Vorschriften der Anlage 3 aber auch dem Stand der Wissenschaft entsprechen. Die in den bisherigen Effizienzvergleichen verwendeten Verfahren zur Dominanz- und Supereffizienzanalyse sind aufgrund methodischer Schwächen nicht hinreichend geeignet, offensichtlich auffällige Unternehmen als Ausreißer zu erkennen. Methodische Anpassungen sowohl bei der Dominanz- und Supereffizienzanalyse sowie der Ausreißeranalyse in der SFA sind notwendig. Nur dieses Vorgehen ist ARegV-konform und stellt zudem eine weitere Annäherung an den Stand von Wissenschaft dar.

Datenbasis muss überprüft werden

Bei der Auswahl der Vergleichsparameter ist die Vergleichbarkeit der teilnehmenden Netzbetreiber sicherzustellen. Neben allgemeinen Standard-Plausibilitätsprüfungen sind die für den Umgang mit Massendaten üblichen Logikprüfungen aller Eingangsdaten durchzuführen, um zu aussagekräftigen Effizienzwerten zu kommen.

Parameterauswahl: Stärken von SFA und DEA nutzen

Gemäß den ARegV-Vorgaben handelt es sich sowohl bei der Stochastic Frontier Analysis (SFA) als auch der Data Envelopment Analysis (DEA) um gleichberechtigte Methoden zur Bestimmung der Effizienz, deren Stärken zur Durchführung eines sachgerechten Effizienzvergleichsverfahrens zwingend genutzt werden sollten. Falls dieses im Ergebnis zu einer unterschiedlichen Parametrierung der Methoden führen sollte, steht dies im Einklang mit den gesetzlichen Vorgaben.

Kostentreiberanalyse muss sachgerecht durchgeführt werden

Eine „optimale Größe“ eines Benchmarking-Modells ergibt sich erst aus dem Umfang und der Struktur des Datensatzes und kann nicht bereits im Vorhinein definiert werden.

Detaillierte Erläuterungen zum Datensatz der Kostentreiberanalyse notwendig

Unregelmäßigkeiten und Auffälligkeiten im verwendeten Datensatz zur Kostentreiberanalyse müssen dargelegt werden. Insbesondere ist eine Lageparameteranalyse notwendig, um ergebnisbeeinflussende Parameter zu identifizieren und auf Plausibilität zu prüfen.

Nachvollziehbarkeit aller Schritte zur Effizienzwertermittlung

Grundsätzlich ist eine hohe Transparenz der zugrundeliegenden Daten und aller Analyseschritte zwingend erforderlich. Alle Schritte zur Modellfindung müssen ausführlich dargestellt und deren Auswirkung mittels Sensitivitätsanalysen dokumentiert werden. Die Regelungen von § 12 ARegV sind zu berücksichtigen: das Benchmarking-Modell der 3. Regulierungsperiode muss konsultiert werden.

Die Kernforderungen werden im Folgenden detailliert dargelegt. Nach einer einleitenden Beurteilung des bisherigen Konsultationsprozesses folgt der Aufbau dann den Schritten zur Modellfindung.

1. Beurteilung des Gesamtprozesses

BDEW/VKU/GEODE haben der BNetzA am 20. Juli 2018 ein Diskussionspapier mit Vorschlägen zum Benchmarking für Stromverteilernetzbetreiber zur Verfügung gestellt. Einzelne Aspekte daraus wurden im Workshop von den BNetzA-Beratern aufgegriffen und diskutiert. Zudem wurden am Tag vor dem Workshop die für die Kostentreiberanalyse verwendeten Stromverteilernetzbetreiberdaten veröffentlicht. Auch waren die im Workshop vorgestellten Inhalte umfangreicher und detaillierter als in der Konsultation zum Benchmarking der Gasverteilernetzbetreiber 2017. Dieses Vorgehen ist mit Blick auf eine offene Diskussion als positiv zu bewerten ebenso wie die ausführlichen Darstellungen der BNetzA-Berater zu einzelnen methodischen Schritten der Modellfindung. Im Ergebnis präsentierten die BNetzA-Berater drei mögliche Benchmarking-Modelle, deren Ausgestaltung allerdings noch nicht final ist.

Die im Folgenden ausgeführten Anmerkungen und offenen Fragen können dennoch nur vergleichsweise allgemein formuliert werden. Insbesondere der Umgang der BNetzA mit dem aus der ARegV-Novelle von 2016 herrührenden Wegfall der Pflichtparameter gemäß § 13 Abs. 4 ARegV a.F. sowie die neu definierte Anforderung an die Berücksichtigung der Heterogenität der Aufgaben der Netzbetreiber gemäß § 13 Abs. 3 Satz 8 ARegV können erst anhand eines tatsächlichen Modells und dem daraus resultierenden Prozess der Effizienzermittlung beurteilt werden. Auf Basis der aktuellen Informationslage ist eine Gesamtbeurteilung eines Modells, der Parameter und des Verfahrens demnach noch nicht möglich.

In dem Termin zur Information der betroffenen Wirtschaftskreise am 25. Juli 2018 wurden drei auch nur als vorläufig bezeichnete Modelle vorgestellt. Soweit ersichtlich, plant die BNetzA derzeit auch nicht, der Branche das von ihr vorgesehene Modell noch zu einem späteren Zeitpunkt vorzustellen und zu konsultieren. Diese Vorgehensweise ist in unseren Augen nicht mit der Regelung gemäß § 12 Abs. 1 ARegV vereinbar, nach der die betroffenen Wirtschaftskreise über die Methoden zu informieren und somit die verwendeten Modelle rechtzeitig zu konsultieren sind. Nach der Begründung des Verordnungsgebers (vgl. BR-Drucks. 417/07, Beschluss, S. 5) soll diese Einbindung der betroffenen Wirtschaftskreise bereits in das Verfahren zur Ausgestaltung des bundesweiten Effizienzvergleiches einerseits die Transparenz und Akzeptanz des Verfahrens erhöhen und andererseits dafür Sorge tragen, dass die spezifischen Kenntnisse der Branche noch Eingang in die finale Ausgestaltung des Effizienzvergleichsmodells finden können.

Im Ergebnis ist daher festzuhalten, dass die **bisherige Konsultation der betroffenen Wirtschaftskreise nur einen ersten Schritt** darstellen kann. Die Vorgaben aus der Verordnung sind jedoch noch nicht vollumfänglich erfüllt. Die BNetzA ist dringend aufgerufen, hier einen weiteren Termin zum Modell der 3. Regulierungsperiode – ggf. auch wie in der Vorperiode erst im Herbst 2018 – stattfinden zu lassen. Nur über eine weitere Einbeziehung der betroffenen Wirtschaftskreise kann der Intention des Verordnungsgebers genügt werden, wie sie den Vorgaben aus § 12 Abs. 1 Satz 2 sowie § 13 Abs. 3 Satz 10 ARegV zugrunde liegen.

Die BNetzA beabsichtigt offenbar, insbesondere Modelle zu untersuchen, die – auch wegen des Wegfalls der bisherigen von der ARegV vorgegebenen Pflichtparameter – mit vergleichs-

weise wenigen Parametern auskommen sollen. Dies steht jedoch klar im Widerspruch zur Vorgabe aus der ARegV, die Heterogenität der Aufgaben der Netzbetreiber abzubilden. Denn durch diese Verpflichtung besteht vielmehr die Notwendigkeit, eine im Vergleich zur zweiten Regulierungsperiode höhere Anzahl (bzw. die Kombination) von Parametern zu untersuchen und mehr Parameter im finalen Modell abzubilden.

Problematisch am bisherigen BNetzA-Vorgehen ist, dass die in **§ 13 Abs. 3 Satz 4 ARegV aufgeführten Parametern** bisher kaum berücksichtigt wurden. Ein Wegfall der sogenannten Pflichtparameter für das Effizienzvergleichsmodell ändert nichts an dem Sachverhalt, dass die vom Ordnungsgeber genannten Strukturparameter, die offensichtlich den Anforderungen gemäß § 13 Abs. 3 Sätze 1-3 ARegV entsprechen, zumindest analysiert und begründet bewertet werden müssen. Denn nur die verpflichtende Aufnahme dieser Parameter ist durch die ARegV-Novelle entfallen. Die aus Sicht des Ordnungsgebers **besondere Bedeutung dieser Strukturparameter** bleibt indes bestehen. Insofern überrascht es, wenn sich die BNetzA und deren Berater von diesen ausdrücklich gesetzlich genannten Parametern im Wesentlichen nur mit denen aus den Ziffern 6 und 7 beschäftigt haben.

Ferner ist derzeit noch **unzureichend auf strukturelle Besonderheiten und die Heterogenität der Aufgaben von Netzbetreibern eingegangen** worden. Zusätzlich ist festzuhalten, dass die Verordnung in § 13 Abs. 3 Satz 1 ARegV als Vergleichsparameter ausdrücklich die Untersuchung von Parametern zur Bestimmung struktureller Besonderheiten der Versorgungsaufgabe aufgrund demografischen Wandels des versorgten Gebietes vorsieht. Soweit ersichtlich, wurde diese Thematik bisher nicht aufgegriffen. Da der Ordnungsgeber diesen Aspekt ausdrücklich in seinem Anforderungskatalog an die Vergleichsparameter aufnimmt, ist es notwendig, dass die BNetzA diesen Prüfungspunkt beachtet.

Die Verordnung sieht ferner vor, dass Vergleichsparameter bezogen auf verschiedene Netzebenen der Stromversorgungsnetze verwendet werden können. Zwar steht eine auf Netzebenen bezogene Verwendung von Vergleichsparametern im Ermessen der Behörde. Bisher ist allerdings nicht bekannt, inwieweit die BNetzA oder die von ihr beauftragten Berater entsprechend Untersuchungen durchgeführt haben. Erst auf Basis von eingehenden Untersuchungen kann eine rechtmäßige Ermessensentscheidung getroffen werden.

Aus alledem ergibt sich, dass das bisherige Verfahren in mehrfacher Hinsicht noch nicht den gesetzlichen Anforderungen, wie sie sich aus § 13 ARegV ergeben, genügt.

Forderungen:

Das Benchmarkingmodell der 3. Regulierungsperiode ist der Branche vor Ermittlung der Effizienzwerte offenzulegen und zu konsultieren, wie es in § 12 ARegV gefordert ist.

Auch wenn die in § 13 Abs. 3 Satz 1 genannten Parameter keine Pflichtparameter mehr darstellen, sollten diese als Ausgangsbasis für die Modellfindung herangezogen werden, d.h. die in § 13 Abs. 3 ARegV aufgeführten Vergleichsparameter müssen im Zuge der Kostentreiberanalyse und Modellfindung eine besondere Würdigung bzw. Berücksichtigung erfahren. Um der Vorgabe gemäß § 13 ARegV zur Abbildung von Heterogenitäten der Aufgaben von Netz-

betreibern nachzukommen, ist es fachlich geboten, Modelle mit einer höheren Anzahl an Vergleichsparametern auszustatten, da diese die Heterogenität der Aufgaben von Netzbetreibern methodisch bedingt erheblich besser abbilden können.

2. Datenqualität und -validierung

Um einen sachgerechten Effizienzvergleich durchzuführen, wird ein **belastbarer Datensatz** benötigt. Die BNetzA hat eine sehr umfangreiche Datenerhebung durchgeführt und diese im Winter 2017/18 sogar noch erweitert. Diese Daten wurden – soweit rechtlich möglich – veröffentlicht, wodurch es der Branche nun erstmalig möglich ist, die Datengrundlage zu untersuchen. Nachfolgende Kapitel enthalten Hinweise zu den Ergebnissen.

2.1 Veröffentlichung und Aufbau des KTA-Datensatzes

Am 24. Juli 2018 veröffentlichte die BNetzA **netzbetreiberindividuelle Aufwands- und Vergleichsparameter** für den Effizienzvergleich der 3. Regulierungsperiode der Stromverteilernetzbetreiber. Die Datei umfasst 204 Stromverteilernetzbetreiber: hiervon sind zum aktuellen Zeitpunkt 26 Datensätze aufgrund laufender gerichtlicher Verfahren geschwärzt. Darüber hinaus liegen für drei Netzbetreiber noch keine Angaben zu den Gesamtkosten (und damit auch nicht zu standardisierten Gesamtkosten) vor. Daneben fehlen insgesamt Angaben zu 70 Variablen in den Bereichen Einspeisung aus EEG/KWKG-Anlagen sowie Ausspeisung an Letztverbraucher. Vor diesem Hintergrund betonte das von der BNetzA mit der Durchführung des Effizienzvergleichs beauftragte Beraterkonsortium im Rahmen der Konsultationsveranstaltung am 25. Juli 2018, dass die Aufwands- und Vergleichsparameter noch nicht abschließend validiert sind (vgl. Sumicsid et al. Folie 39). Erst mit einem vollständigen und validierten Datensatz ist eine Modellfindung möglich.

Die Datei umfasst neben den Aufwands- und Vergleichsparametern auch eine Überführung der KTA-Variablenamen in den von den Stromverteilernetzbetreibern befüllten Strukturdatenerhebungsbögen bzw. Zusatzerhebungsbögen sowie eine Erläuterung der Berechnung der daraus gebildeten Vergleichsparameter. Diese Vereinfachung der Plausibilisierungsmöglichkeiten wird durch die Branche begrüßt.

Zwei Punkte sind in der **KTA-Datei zu ergänzen bzw. zu korrigieren**:

- a) Die Parameter der Gebietseigenschaften (Nr. 647 bis 652) lassen sich nicht nachvollziehen. Hier ist die Bildung aus Leitungslänge und Bodenklassen- bzw. Grabbarkeitsflächenanteilen zu ergänzen.
- b) Im Bereich der Ausspeisung aus Kaskadierung ist irrtümlich der Parameter Nr. 528 („yEnergy.recovered.from.higher.net.hoes“) ausgewiesen, obwohl diese Größe technisch nicht existiert und im Erhebungsbogen auch nicht abgefragt wird. Die hierfür in der KTA-Datei genannten Werte entsprechen dem Parameter Nr. 529 („yEnergy.recovered.from.higher.net.hoes_hs“), der somit doppelt auftritt.

Die BNetzA sieht im weiteren Vorgehen vor, die Datenerhebung sowie die Konsultation der bisherigen Parameterauswahl bis zum 16. August 2018 abzuschließen und Resultate der danach erfolgenden Berechnungen und Validierungen im September vorzustellen. Endgültig kann jedoch derzeit nicht beurteilt werden, wann die Datenerhebung abgeschlossen sein wird. Vor diesem Hintergrund sowie der im Folgenden dargestellten Auffälligkeiten ist aus Branchensicht eine regelmäßige Aktualisierung des veröffentlichten KTA-Datensatzes, mindestens aber die Veröffentlichung des finalen Standes, notwendig.

Forderungen:

Eine regelmäßige Aktualisierung und Veröffentlichung des bisher unvollständigen KTA-Datensatzes ist notwendig.

Der finale Stand des KTA-Datensatzes ist zu veröffentlichen.

Die den Gebietseigenschaften (Zeile 647 bis 652, Spalte A) zugrundeliegende mathematische Logik ist analog zu den anderen Parametern in Spalte H darzulegen.

Zur Erhöhung der Transparenz sollten die Netzbetreiber über die durchgeführten Anpassungen in elektronischer Form informiert werden.

2.2 Behördliche Plausibilisierungsprüfungen

In der BNetzA-Unterlage „Konsultation zur Auswahl der Vergleichsparameter zum Effizienzvergleich der deutschen Elektrizitätsverteilernetzbetreiber“ vom 25. Juli 2018 wird angegeben, dass umfangreiche **netzbetreiberindividuelle Plausibilisierungen** vorgenommen wurden. Diese beinhalten laut Foliensatz unter anderem Logikprüfungen und Vergleiche mit bereits gelieferten Daten in anderen Verfahren und Altdaten. Bezüglich Letzteren ist jedoch nicht bekannt, welche Daten verwendet wurden und welche Kriterien bei der Bewertung angesetzt wurden. Neben einer Vielzahl an Vollständigkeitsprüfungen, unter anderem in Form bedingter Formatierungen im Erhebungsbogen erfolgte zusätzlich die Bereitstellung eines Word-Dokuments², welches die von der Behörde durchgeführten Logikprüfungen mit technischem Bezug beinhaltet. Diese Prüfungen sind aus unserer Sicht sowohl hinsichtlich der einbezogenen Daten als auch der **Tiefe der Prüfungslogiken** nicht ausreichend. Eine stichprobenartige Analyse lässt in Bezug auf die Streuung im Datensatz vermuten, dass unter anderem in der Datengruppe „Länge der Hausanschlussleitungen“ unübliche und unterschiedliche Verhältnisse zu Anschlusspunkten der Niederspannung bestehen, ebenso bei „Anzahl der Zähler“ im Verhältnis zur Gesamtbevölkerung. Vor diesem Hintergrund sind tiefgreifende individuelle ingenieurwissenschaftliche Prüfungen notwendig, um die durchgehende Qualität der Angaben sowie deren Vergleichbarkeit zu gewährleisten.

Darüber hinaus ist es unter dem Aspekt der Transparenz und Nachvollziehbarkeit unerlässlich, dass die BNetzA eine konkrete Beschreibung der Vorgehensweise zur Datenplausibilisierung publiziert – insbesondere die verwendeten methodischen Konzepte, die Wertebereiche, die zur Prüfung von Kennzahlen verwendet wurden und der Umgang mit Extrempositionen bei

² BNetzA (März 2018), Beschreibung der Prüfungen

Kostenkennzahlen, die auf mögliche Alleinstellungsmerkmale, insbesondere in der DEA, hindeuten können.

Laut BNetzA-Aussagen bei dem Workshop wurden außerdem gesonderte Prüfungen der geografischen und versorgten Flächen sowie zu den Angaben bzgl. der Einspeisemanagement- und Investitionsmaßnahmen durchgeführt. Zusätzlich erfolgte stichprobenhaft ein Abgleich mit öffentlich verfügbaren Daten. Notwendig ist eine Erläuterung, welche öffentlich verfügbaren Daten zugrunde gelegt wurden und wie sichergestellt wurde, dass diese korrekt sind. Darüber hinaus muss dargelegt werden, ob und in welchem Umfang Stromverteilernetzbetreiber von einer individuellen Sonderprüfung (Folie 31) betroffen waren.

Forderungen:

Die Einführung von Logikprüfungen, die auf ingenieurwissenschaftlichen Zusammenhängen beruhen bzw. definitionsgleiche Datenbestände anderer Abfragen beinhalten, ist notwendig.

Die mathematischen Formeln sind zur Erhöhung der Transparenz bei allen durchgeführten Logikprüfungen und somit auch für die Gebietseigenschaften darzulegen.

2.3 Analyse der Datengrundlage für die in der Konsultation vorgestellten Parameter

Als Referenzwert für einen Vergleichsparameter werden die besten (d.h. niedrigsten) Stückkosten herangezogen, wobei auf Basis der Erläuterungen im Rahmen der Konsultationsveranstaltung viele der abgefragten Parameter nicht in die finale Berechnung des Effizienzwerts eingehen werden. Aus diesem Grund wurde im BDEW/VKU/GEODE-Projekt „BMT“ zunächst eine gesonderte Untersuchung aller Parameter vorgenommen, die in den aktuell von Swiss economics, Sumicsid und IAEW favorisierten drei Grundmodellen zur Anwendung kommen. Dabei wurde deren Unterscheidung in kleine und große Unternehmen zunächst vernachlässigt. Die Grundmodelle umfassen folgende Parameter:

RP2	BU1	TD2
Anz. Anschluss- und Einspeisepunkte	Summe Ausspeisung gesamt	Zeitgl. Jahreshöchstlast HöS/HS & HS/MS
Stromkreislänge Kabel HS	Anzahl Zählpunkte im Netzgebiet (ohne Pauschalanlagen)	Anzahl selbst betriebener Zählpunkte (mit Pauschalanlagen)
Stromkreislänge Kabel MS	Stromkreislänge Freileitung gesamt	Stromkreislänge Freileitung Total inkl. Straßenbeleuchtung
Stromkreislänge Freileitung HS	Installierte Bemessungsscheinleistung	Inst. Erzeugungsleistung Erneuerbare gesamt
Stromkreislänge Freileitung MS	Installierte Erzeugungsleistung Erneuerbare Wind	
Stromkreislänge NS		
Zeitgl. Jahreshöchstlast HS/MS		
Zeitgl. Jahreshöchstlast MS/NS		
Installierte Erzeugungsleistung Erneuerbare gesamt		
Versorgte Fläche NS		

Anzahl selbst betriebener Zähl- punkte (mit Pauschalanlagen)		
---	--	--

Für diese 20 Parameter wurden im Rahmen des BMT-Projektes der Verbände die Angaben der Unternehmen in der **KTA-Datei auf mögliche Unregelmäßigkeiten** untersucht. Im Rahmen einer Stichprobe von Angaben im veröffentlichten Datensatz sind folgende Punkte aufgefallen:

- **Industrieanlagen:** Es handelt sich mehrfach um „Netzbetreiber“, welche die Versorgung der eigenen Produktion mit infrastrukturellen Dienstleistern sichern. Daraus folgt stets eine sehr geringe Anzahl von Zählpunkten bei gleichzeitig hohen übertragenen Energiemengen. Darüber hinaus betreiben diese nur eine oder zwei Spannungsebenen und es erfolgt – im Unterschied zu „regulären“ Stromverteilernetzbetreibern – keine Einspeisung aus dezentralen, regenerativen Energiequellen.
- **Eingabefehler:** Zwei Netze weisen einen Anteil von Zählpunkten mit Leerstand von exakt 50 % auf, was auf einen Übertragungsfehler hindeutet.

Eine über diese 20 Parameter hinausgehende deskriptive Analyse zeigt, dass sich einige wenige Netzbetreiber in Struktur und/oder Größe erheblich von der Masse der Stromverteilernetzbetreiber unterscheiden. Beispiele hierfür sind:

- Leitungslänge pro km² versorgter Fläche: Abweichungen im Bereich -100% bis +200% zum Median
- Industrienetze: Verhältnis von Bevölkerungszahl zu Jahreshöchstlast (JHL) gleich null, extreme Abweichungen bei JHL pro Anschlusspunkt (> 10.000 % im Median) im Vergleich zu anderen Verteilernetzbetreibern
- Dezentrale installierte Leistung zu JHL: Schwankungen um den Mittelwert (95%) in einer Spanne von 0 % - 400 %
- Einspeisepunkte: Netzbetreiber mit nur einem Einspeisepunkt von Erzeugungsanlagen weisen extreme Abweichungen im Verhältnis zur zeitgleichen JHL im Vergleich zu anderen Verteilernetzbetreibern aus.

Forderung:

Vor dem Hintergrund dieser Besonderheiten sollte die BNetzA die KTA-Datei hinsichtlich struktureller Unterschiede zwischen den Stromverteilernetzbetreibern untersuchen und Fehler bzw. Auffälligkeiten korrigieren sowie die Ergebnisse hierzu veröffentlichen.

2.4 Plausibilisierung des KTA-Datensatzes durch Lageparameteranalyse

Zur Analyse der Parameter, die potentiell für einen Effizienzvergleich herangezogen werden (**KTA-Datensatz**), wurden über alle Netzbetreiber hinweg für jede Variable das arithmetische Mittel, der Median sowie die durchschnittliche Abweichung von beiden Mittelwerten gebildet. Darüber hinaus dienen das 1., 5., 95. und 99. Perzentil als Orientierung für die Verteilung der Werte über die Stichprobe hinweg. Eine Untersuchung der Netzbetreiber, die besonders häu-

fig über dem 95. Perzentil oder unter dem 5. Perzentil lagen, bestätigte die Ergebnisse, die bereits durch die Verhältniskennzahlen deutlich wurden: Werte unter dem 5. Perzentil treten sehr verstreut auf und wirken wenig systematisch geprägt.

Ein ganz anderes Bild liefern die Werte über dem 95. Perzentil (Abbildung 1): Die besonders hohen Parameter-Werte konzentrieren sich hier auf einige wenige Netzbetreiber, die in bis zu 345 Variablen zu den höchsten Werten zählen. Dies spiegelt die erheblichen Größenunterschiede am Markt wider und unterstreicht die Heterogenität der Aufgaben von Netzbetreibern in der Stichprobe.

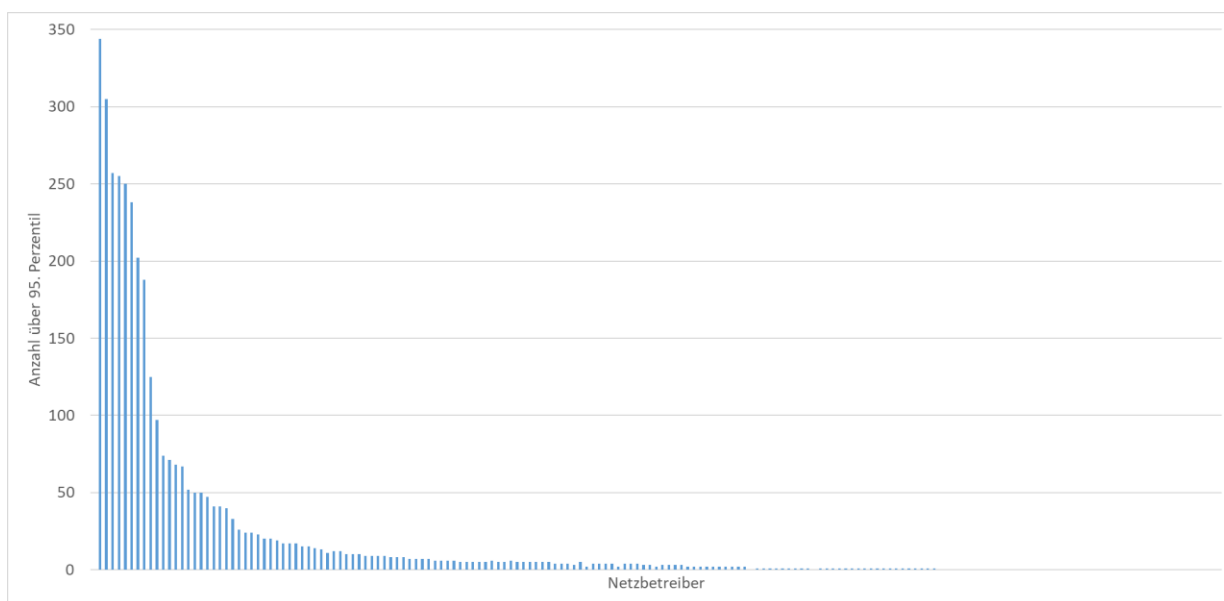


Abbildung 1: Anzahl der Überschreitungen des 95. Perzentils in den einzelnen Parameterwerten

Für eine von der Größe des Netzbetreibers unabhängige Untersuchung der veröffentlichten Daten wurde zusätzlich die Struktur der Verhältnisse zwischen den Gesamtkosten eines Netzbetreibers und deren Parameterwerten betrachtet. Die Struktur der "Unit Costs", die anhand des Verhältnisses zwischen der standardisierten und nicht-standardisierten Gesamtkosten (TOTEX bzw. sTOTEX) zu den einzelnen abgefragten Parametern (exkl. derjenigen abgefragten Variablen, die nur die Werte 0 oder 1 annehmen können) errechnet werden, zeigt den erheblichen Einfluss einiger strukturell stark vom Durchschnitt abweichender Netzbetreiber. Abbildung 2 verdeutlicht in Anlehnung an Folie 38 des zur Konsultation bereitgestellten Foliensatzes, wie häufig ein Netzbetreiber das beste (d.h. niedrigste) Kosten/Parameter-Verhältnis aufweist. Die Ergebnisse sind nach Anzahl der Bestergebnisse von links nach rechts sortiert.

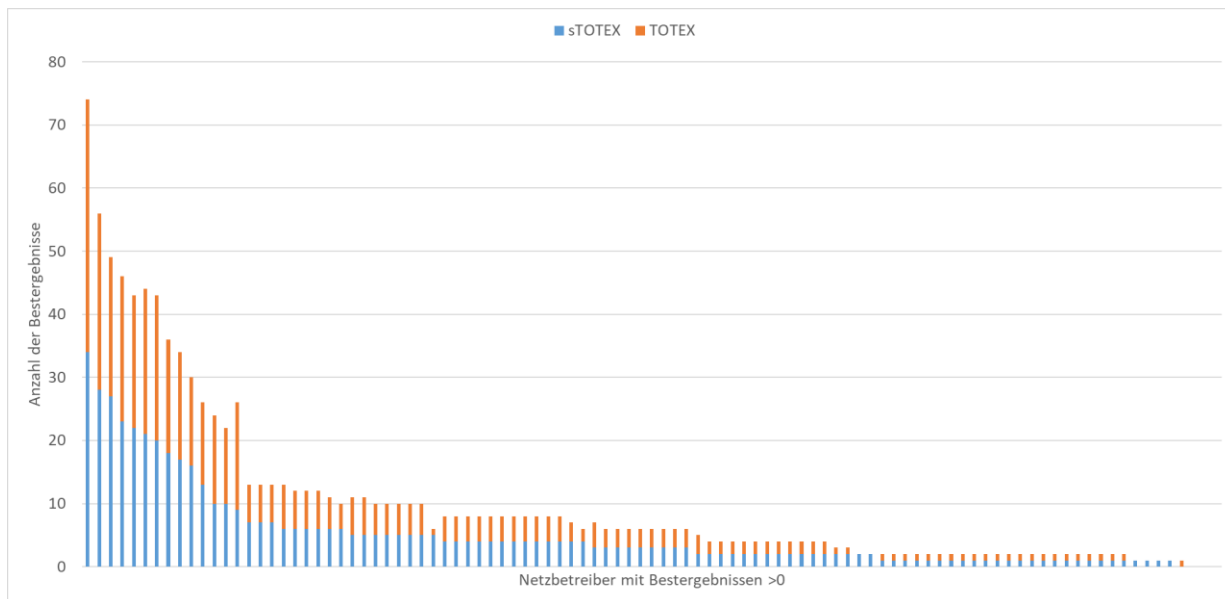


Abbildung 2: Netzbetreiber mit einem oder mehr Bestwerten im Ranking der Unit Costs, sortiert nach Anzahl der Bestwerte in sTOTEX-basierter Rechnung (exkl. geschwätzter Variablen und Variablen, die mit Ja/Nein beantwortet wurden)

Die 14 Netzbetreiber mit den häufigsten Bestergebnissen fallen besonders dadurch auf, dass diese entweder nur sehr wenige Spannungsebenen bedienen und insgesamt sehr klein sind (vgl. dazu auch die Definition von „kleinen“ und „großen“ Netzbetreibern durch die BNetzA, Folie 38 des BNetzA-Foliensatzes), oder aber besonders große Netzbetreiber. Dies wird noch deutlicher, wenn die oben für alle Netzbetreiber gezeigte Grafik nach Größe des Netzbetreibers, gemessen in der absoluten Höhe der (standardisierten) Gesamtkosten unterteilt wird. Abbildung 3 zeigt die Ergebnisse von oben, absteigend unterteilt nach 10 Größenklassen, die durch die jeweiligen Dezile der standardisierten TOTEX unterteilt werden.

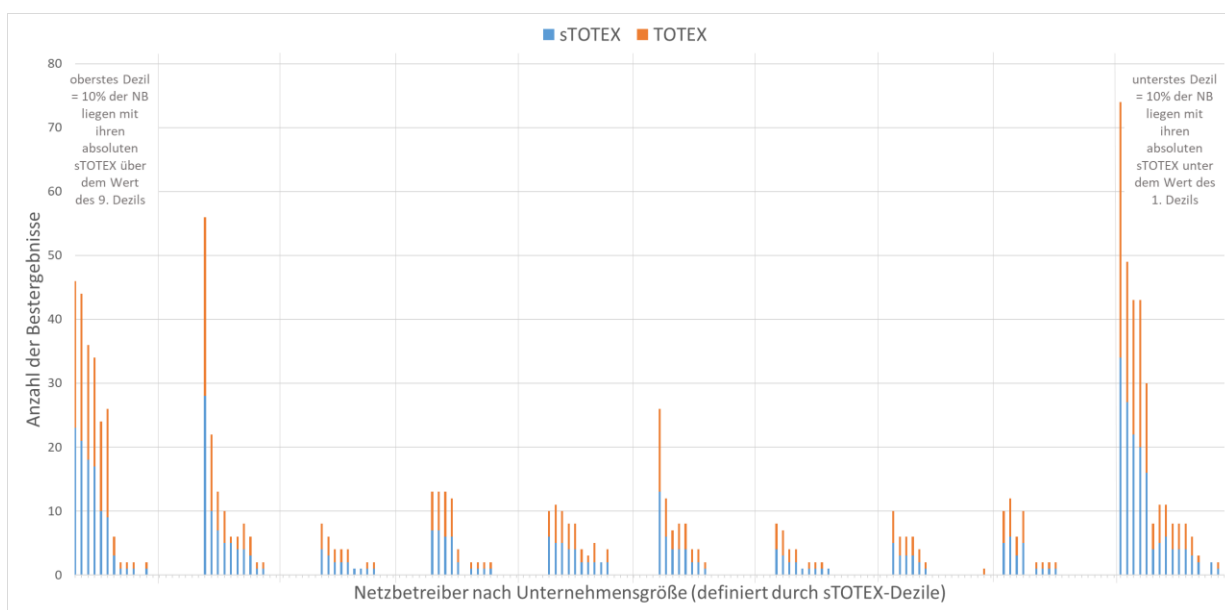


Abbildung 3: Netzbetreiber nach Bestwerten im Ranking der Unit Costs nach Anzahl der Bestwerte in sTOTEX-basierter Rechnung und Dezil der sTOTEX-Höhe (exkl. geschwätzter Variablen und Variablen, die mit Ja/Nein beantwortet wurden)

Die BNetzA-Berater erläuterten bei dem Workshop, dass Stromverteilernetzbetreiber mit besonders vielen Bestergebnissen vermehrt als „kleine“ oder „große“ Netzbetreiber identifiziert wurden. Dieses wird auch aus Abbildung 3 ersichtlich: bei besonders „großen“ (vgl. Folie 38 sumicisid et al.) Unternehmen (ganz links werden Netzbetreiber gezeigt, die größer als das neunte Dezil der sTOTEX-Verteilung sind, also sTOTEX von mehr als rd. 169 Mio. Euro aufweisen) oder aber bei „kleinen“ Netzbetreiber (vgl. Sumicisid et al. Folie 38), wie die Verteilungen ganz rechts zeigen.³

„Kleine“ Netzbetreiber mit Bestwerten bedienen bzw. betreiben (fast) ausnahmslos wenige oder eine einzige Spannungsebene. Eine Vielzahl der Verteilernetzbetreiber dagegen betreiben generell mehrere Spannungsebenen. Die BNetzA sollte – wie auch angekündigt – prüfen, ob hier strukturelle Unterschiede in der Versorgungsaufgabe vorliegen und in welcher Form dieses ggf. zu berücksichtigen ist bei der **Ermittlung von sachgerechten Effizienzwerten für alle Stromverteilernetzbetreiber**.

Erklärungsbedürftig ist auch die Vielzahl an Stromverteilernetzbetreibern mit nur einem Alleinstellungsmerkmal in den Parametern. Dies kann entweder auf Datenfehler hindeuten oder auf Alleinstellungsmerkmale, d.h. der Netzbetreiber ist in seiner Versorgungsaufgabe nicht mit anderen Netzbetreibern vergleichbar. Diese Alleinstellungsmerkmale sind ggfs. gesondert zu überprüfen und die Auswirkungen auf die DEA zu analysieren (vgl. Sumicisid et al. Folie 39).

Hinzu kommen zwei besonders auffällige Verteilernetzbetreiber in den Dezilen 2 und 6, die in ihrer Größenklasse ungewöhnliche Kostenstrukturen aufweisen. Dies betrifft zwei Netzbetreiber, die innerhalb der (nach Dezilen der Unternehmensgröße unterteilten) Vergleichsgruppe besonders viele Bestergebnisse erzielen, wie in Abbildung 3 zu erkennen ist.

Forderungen:

Die BNetzA sollte strukturelle Unterschiede in der Versorgungsaufgabe (insbesondere die immensen Größenunterschiede) der Stromverteilernetzbetreiber auf allen Stufen des Vorgehens berücksichtigen.

Zudem ist der angestrebte Umgang im Rahmen des Effizienzbenchmarkings der 3. Regulierungsperiode mit strukturell nicht vergleichbaren Netzbetreibern zu erläutern.

Der Umgang mit den Größenunterschieden der Stromverteilernetzbetreiber im Effizienzvergleich ist transparent darzulegen.

Daten-Auffälligkeiten hinsichtlich struktureller Unterschiede sollten von der BNetzA untersucht, gegebenenfalls korrigiert und die Ergebnisse veröffentlicht werden.

³ Laut Definition der BNetzA ist ein „kleiner“ Netzbetreiber durch Totex < 5 Mio. Euro definiert, ein „großer“ Netzbetreiber durch sTOTEX > 100 Mio. Euro. Die Unterteilung der Unternehmen in Dezile nach sTOTEX entsprechen dieser Definition weitestgehend: Das erste Dezil entspricht 5,98 Mio. Euro, das zehnte Dezil einem Wert von 168,84 Mio. Euro.

2.5 Transparenz der Datengrundlagen

Die Stromverteilernetzbetreiber müssen neben den von der BNetzA bereits zur Verfügung gestellten Strukturdatenquittungen auch eine Kostenquittung von ihrer jeweiligen Regulierungsbehörde erhalten, die als Excel-Datei die in den Effizienzvergleich eingehenden Kostendaten enthält. Von den Landesregulierungsbehörden Bayern, Baden-Württemberg, Hessen und Rheinland-Pfalz liegen solche Kostenquittungen mit dem Titel „Vergleichbarkeitsrechnung für Effizienzvergleich“ bereits vor, aus der (a) die konsolidierten Netzkosten und die dauerhaft nicht beeinflussbaren Kostenanteile, jeweils nach Kostenarten gegliedert, und die daraus abgeleiteten Aufwandsparameter mit genehmigten und standardisierten Kosten hervorgehen und die (b) auch die Vergleichbarkeitsrechnung des konsolidierten Anlagevermögens einschließlich der Berechnung der annuitätischen Kosten und der zusätzlichen Zinsen enthält. Auch die BNetzA und die übrigen Regulierungsbehörden sollten den Netzbetreibern in ihrer Zuständigkeit diese **Kostenquittungen zur Verfügung stellen**.

Neben einer ausführlichen und verständlichen Dokumentation der Ergebnisse in einem Ergebnisbericht sollte die BNetzA „do-files“ und „log-files“ in hinreichend dokumentierter Form zur Verfügung stellen. Aus diesen sollte hervorgehen, welche Befehle im Schätzprogramm ausgeführt wurden und welche Qualität die jeweils erzielten Ergebnisse aufweisen. Ein Ausweis von Einzeldaten oder -ergebnissen soll nicht erfolgen.

Forderungen:

Änderungen des Datensatzes sind regelmäßig zu veröffentlichen und den Netzbetreibern in elektronischer Form unverzüglich mitzuteilen.

Die BNetzA sollte den Netzbetreibern zur Nachvollziehbarkeit der Datenverwendung folgende Dokumentationen zur Verfügung stellen:

- eine vollständige Dokumentation der Ergebnisse
- eine Dokumentation der „do-files“ und „log-files“.

3 Kostentreiberanalyse und Effizienzmessung

3.1 Starker Fokus auf statistische Kostentreiberanalyse

Mit der **statistischen Kostentreiberanalyse** soll eine Vorauswahl möglicher Modelle gefunden werden, die dann sukzessive (händisch) weiterentwickelt wird. Bei der Vorgehensweise zur Identifikation des Benchmarkingmodells und der Ermittlung der Effizienzwerte wählen die BNetzA-Berater eine Vorgehensweise, die in Bezug auf wissenschaftliche Kriterien anzuzweifeln ist. Zu erwähnen ist beispielsweise die Verwendung einer linearen Funktion auf Niveaudaten für die Kostentreiberanalyse, die Herleitung der „optimalen Modellgröße“, die starke Betonung der Vermeidung von Multikollinearität, der ungenügende Umgang mit der Frage der Pfadabhängigkeit und damit verbunden der fehlende Einsatz neuer Methoden zur Kostentreiberanalyse. Im Folgenden werden diese einzelnen Punkte konkretisiert.

3.1.1 Annahme der linearen Funktion

Die BNetzA-Berater führen die **Kostentreiberanalyse auf einem linearen Modell** mit Niveaudaten durch. Die Gründe für dieses Vorgehen bleiben auch nach der Konsultation unklar. So ist insbesondere nicht nachvollziehbar, weshalb die Berater auf entsprechende Fragen nach einer logarithmierten Funktion argumentieren, dies entspreche nicht einer Cobb-Douglas Kostenfunktion. Ebenso unklar ist die Aussage, dass log lineare Modelle „Probleme mit der Biegung (curvature)“ haben (Sumicsid et al. Folie 97). Ein entsprechender Beleg dieser Aussage aus der wissenschaftlichen Literatur (z. B. Simulationsrechnung) ist nicht erfolgt. Das Vorgehen der BNetzA-Berater ist vor allem vor dem Hintergrund der Probleme der Heteroskedastizität aufgrund der extremen Größenunterschiede der untersuchten Verteilernetzbetreiber problematisch.

Lineare Regressionen, welche zur **Modellfindung** herangezogen werden, basieren auf einem Set von Annahmen. Insbesondere die Verletzung der Annahme der Homoskedastizität der Störterme kann sowohl die Modellfindung auf OLS-Basis als auch die Effizienzermittlung in der SFA beeinflussen. Bei Vorliegen von Heteroskedastizität sind bei Anwendung von OLS-Regressionen im Rahmen der Kostentreiberanalyse die Varianzen der OLS-Koeffizienten verzerrt. Es resultieren falsche Ergebnissen in Hypothesentests und verzerrte Ergebnisse in den gemäß den BNetzA-Beratern anzuwendenden t-Tests (Sumicsid et al. unnummerierte Folie vor Folie 63). **Heteroskedastizität** muss in diesem Schritt als wichtiger Hinweis auf möglicherweise fehlende relevante Variablen oder eine inkorrekt spezifizierte funktionale Form angesehen werden (vgl. Kennedy, 2008, S. 126). Erst nach einer Überprüfung der funktionalen Form und dem Ausschluss des Vorliegens eines „omitted variable bias“ (statistisch sowie anhand von energiewirtschaftlichen und ingenieurwissenschaftlichen Überlegungen) sollten die Kostentreiberanalyse basierend auf der finalen funktionalen Form durchgeführt bzw. überprüft werden.

Das Verzerrungspotential kann auf den bestehenden von der BNetzA publizierten Daten aufgezeigt werden. Werden im Sinne einer „Bottom-Up“ Untersuchung 2.297 verschiedene Modelle gerechnet, existierten lediglich zwei (bzw. 79) Modelle, bei denen auf Basis des White (bzw. Breusch-Pagan) Tests die Nullhypothese (konstante Varianz der Residuen) auf einem 5% Konfidenzniveau nicht abgelehnt werden können. Auch die hohen R^2 der Modelle, welche auf der Folie 94 des Foliensatzes der BNetzA-Berater präsentiert wurden, zeigen deutlich, dass die Varianz der Kosten sehr groß ist und dass große Kosten mit anderen großen Variablen korrelieren. Es ist zu erwarten, dass diese ausgeprägte Heterogenität bezüglich der Netzgröße nicht nur zu kleineren Verzerrungen der Standardfehler führt (was man mit robusten Standardfehlern noch heilen könnte), sondern auch, dass die Schätzkoeffizienten verzerrt sein dürften. Auf Basis dieser linearen Regressionen kann daher keinerlei Aussage zu möglichen Vergleichsparametern getroffen werden. Stattdessen sollte die Kostentreiberanalyse auf den gleichen Modelltypen durchgeführt werden, die am Ende auch für die Effizienzmessung verwendet werden (also normiert-lineare Modelle oder Log/log-Modelle).

Dabei zeigt sich, dass die **unterschiedlichen Modelltypen unterschiedliche Kostentreiberkombinationen** als mögliche Modellkandidaten identifizieren. Dies verdeutlicht Abbildung 4,

welche darstellt, wie häufig eine Variable in den verschiedenen Modellen zur „Bottom-Up“ Untersuchung signifikant erschien. Die Variablennummerierung entspricht dabei der Datenveröffentlichung. Zum Beispiel wird die Variable 190 (rot markiert) in dem linearen Modell relativ häufig als signifikant erkannt, spielt aber für die Mehrzahl der normiert-linearen oder Log/log-linearen Modellen keine Rolle. Umgekehrt ist Variable 170 (blau markiert) in normiert-linearen Modellen eher relevant als in linearen Modellen.

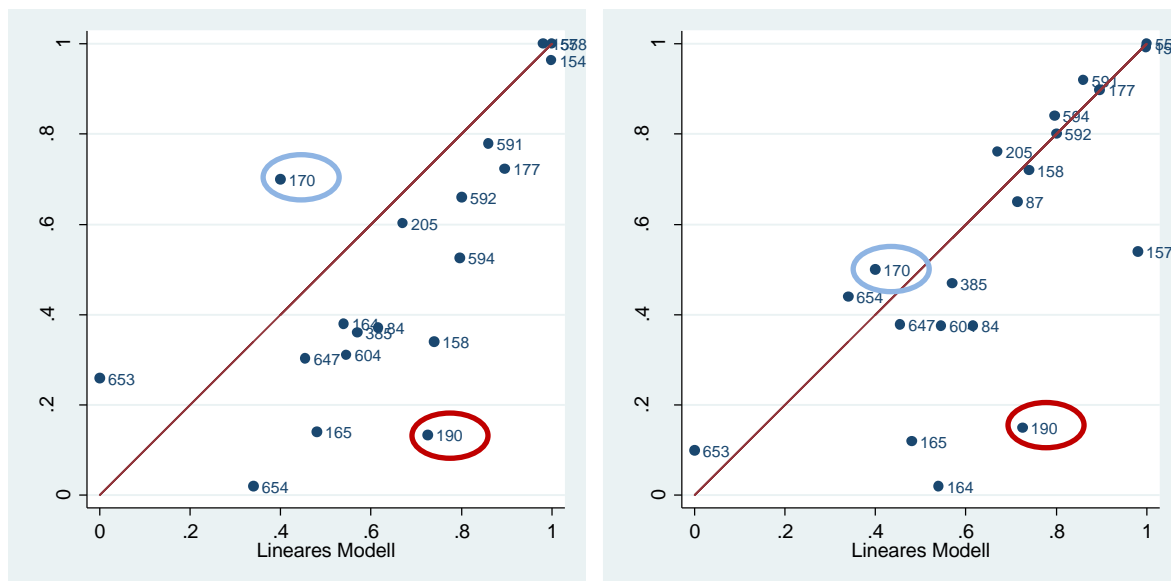


Abbildung 4: Relevanz der verschiedenen Vergleichsparameter

Forderungen

Die Kostentreiberanalyse ist zwingend unter Berücksichtigung der funktionalen Form durchzuführen, die auch für die Bestimmung der Effizienzwerte angewendet wird, wobei verschiedene Grundmodelle auszutesten sind.

Ein Abstellen auf lineare Modelle ist aufgrund der großen Heterogenität in den Daten nicht angebracht und führt zu einer falschen Identifikation von Kostentreibern.

Aufgrund der Tatsache, dass bei Vorliegen von Heteroskedastizität die Varianzen der geschätzten Koeffizienten in der Kostentreiberanalyse verzerrt sind – und damit zu einer falschen Auswahl von Vergleichsparametern führen kann – ist diese zwingend zu korrigieren.

Auf der Stufe der SFA-Effizienzmessungen ist zusätzliche Vorsicht zur Sicherstellung der Homoskedastizität notwendig, weil ansonsten die Effizienzwerte verzerrt sind.

3.1.2 Bestimmung der „optimalen Größe“ eines Modells

Die BNetzA-Berater legen bei der **Modellfindung** ein starkes Gewicht auf die „**optimale Modellgröße**“. So zeigen sie auf Folie 45 des Foliensatzes von Swiss Economics und Sumicsid exemplarisch, dass bei zu wenigen Parametern ein „approximation error“ und bei einem „zu großen“ Modell mit „zu vielen“ Parametern ein „estimation error“ entsteht. Die BNetzA-Berater

präsentierten hinzukommend ein Ergebnis, welches belegen soll, dass die „optimale“ Anzahl an Parametern bei zirka sechs bis acht Parameter liegt und erklären, dass bei mehr als acht Parametern der BIC wieder ansteige. Weiter führt der BNetzA-Berater aus, dass Modelle mit mehr Parametern den Erklärungsgehalt (gemessen am Adj R²) nicht mehr erhöhen würden. Die BNetzA-Berater fokussierten bei ihrer statischen Beurteilung der Modellgüte und bei der Beurteilung der Anzahl Parameter sehr stark auf das Gütemaße BIC (Bayesian information criterion). Aus Robustheitsgründen sind grundsätzlich auch weitere Gütemaße heranzuziehen.

Zudem sind die aus der Abbildung auf Folie 70 gezogenen Schlussfolgerungen der BNetzA-Berater irreführend. Die Abbildung zeigt, dass der **Erklärungsgehalt eines Modells** mit zwei oder drei Parametern bereits bei 98.6% (Adj R²) liegt. Die Erhöhung des Erklärungsgehalts mit 11 Parametern liegt mit 99.8% lediglich gut einen Prozentpunkt höher. Der hohe Erklärungsgehalt bereits mit zwei oder drei Modellparametern zeigt auf, dass aufgrund der gewählten Vorgehensweise der BNetzA-Berater, die Kostentreiberanalyse auf Niveaudaten durchzuführen (vgl. Abschnitt 3.1.1), primär die Größenunterschiede der Netzbetreiber erklärt werden (hohe Kosten korrelieren mit langen Netzen!).

Ebenso wenig aussagekräftig ist die auf Folie 71 dargestellte Analyse des BIC, um die Aussage im Titel zur Folie formulierte Aussage zu treffen, dass das „BIC bei ca. sechs Parametern optimiert“ wird. Zum einen kann diese Aussage nicht überprüft werden, da die Abbildung bei Modellen mit acht Parametern endet. Die Aussage der Berater, dass das BIC nach Verwendung von Modellen mit mehr als acht Parametern wieder steigt, ist unbelegt. Zum anderen ist nicht dokumentiert, wie viele Modelle mit jeweils der betrachteten Anzahl Parameter in die Analyse eingeflossen sind und wie diese Modelle ausgewählt wurden. Aus diesem Grund wurden eigene „Bottom-Up“-Analysen im Rahmen des BMT-Projektes durchgeführt. Die Analysen erfolgten dabei sowohl auf Niveaudaten (wie dies von den BNetzA-Beratern durchgeführt wurde), als auch auf normiert-linearen und log/log linearen Modellen. Letzteres aus dem Grund, da lineare Modelle mit Niveaudaten nicht geeignet sind, um Kostentreiberanalysen durchzuführen (vgl. Abschnitt 3.1.1).

Es ist zu beachten, dass bei dieser „Bottom-Up“-Untersuchung eine Reihe von Modellen zu einer sehr großen Anzahl von Ausreißern führt (basierend auf hohen Cook's Distance Werten). Gütemaße wie der BIC sind abhängig von der Größe der verwendeten Stichprobe. Modelle mit sehr niedrigen BIC (lineare Modelle) sind nicht zwangsläufig besser, sondern wurden lediglich auf einer kleineren Stichprobe berechnet. Eine getrennte Abbildung der BIC für verschiedene Modelle ist zudem notwendig, da die abhängige Variable unter jeder Spezifikation unterschiedlich ist und die BIC-Werte daher nicht direkt miteinander vergleichbar sind.

Abbildung 5 zeigt das Ergebnis, dass Modelle mit sechs Parametern nicht generell besser sind als Modelle mit mehr Parametern. Der geringste BIC-Wert verbleibt auf vergleichsweise ähnlichem Niveau und steigt nach acht Vergleichsparametern nicht wieder an. Allerdings zeigt die vollständige Darstellung der Ergebnisse, dass das Risiko eines statistisch gesehen „schlechten“ Modells mit der Anzahl von Vergleichsparametern sinkt, denn die Streuung der BIC-Werte nimmt deutlich ab. Es kann auch statistisch gesehen „gute“ Modelle mit weniger Parametern

geben. Vollständigere Modelle reduzieren allerdings das Risiko, dass relevante Vergleichsparameter bei einer zu restriktiven Parameterauswahl nicht aufgenommen werden:

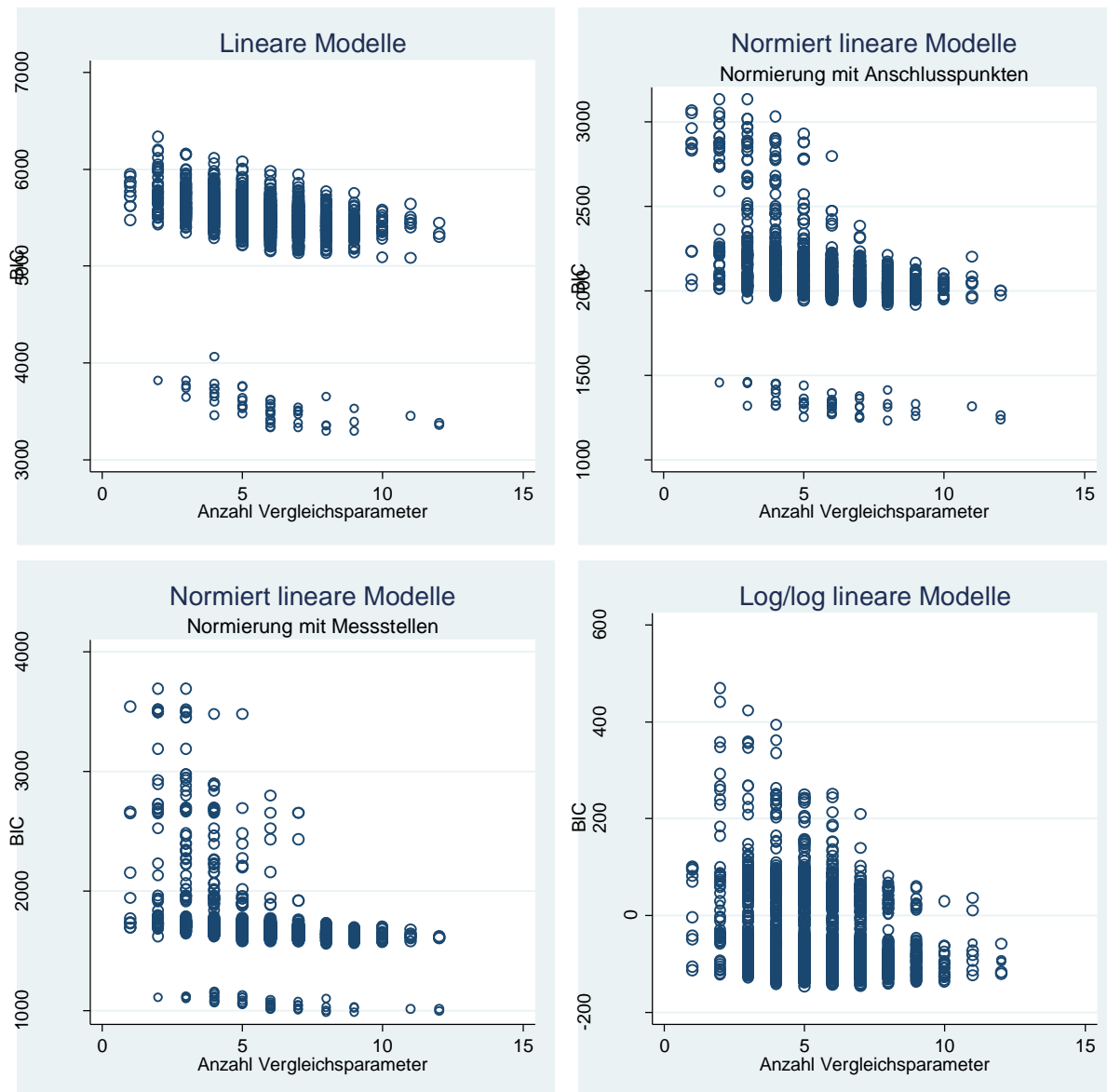


Abbildung 5: Vergleich BIC mit Anzahl der Vergleichsparameter

Forderungen

Eine Fokussierung auf sechs bis acht Parametern bei der Modellfindung ist nicht sachgerecht, da Analysen zeigen, dass auch Modelle mit mehr Parametern gute statistische Eigenschaften besitzen können. Umfang und Struktur des Datensatzes sind entscheidend für eine sachgerechte Größe des Modells.

Größere Modelle reduzieren zudem das Risiko, dass relevante Vergleichsparameter nicht aufgenommen werden. Größere Modelle entstehen insbesondere auch, wenn auf eine (aus Sicht der Kostentreiberanalyse ohnehin fragliche) Aggregation von Einzelstrukturdaten verzichtet wird.

3.1.3 Zu starke Gewichtung auf die Problematik der Multikollinearität

Die BNetzA-Berater betonen an verschiedenen Stellen (z. B. nicht-nummerierte Folie im Kapitel 4.2), dass bei der Beurteilung der Modellgüte und der Parameterwahl keine **Multikollinearität** vorherrschen sollte. So ist für die Berater beispielsweise das „falsche“ Vorzeichen eines Parameters ein Grund, diesen im weiteren Modellfindungsprozess zu ersetzen. „Falsche“ Vorzeichen sind dabei oft das Ergebnis einer Kollinearität dieses Parameters mit anderen im Modell berücksichtigten Parametern und stellen bei der Effizienzmessung kein Problem dar. Die starke Betonung der Signifikanz eines Parameters und der Multikollinearität ist im Kontext der Effizienzmessung nicht angebracht. Denn Multikollinearität stellt in den Modellen aus statistischer Sicht kein Problem dar, wenn keine Hypothesentests für einzelne Variablen durchgeführt oder keine Voraussagen bezüglich einzelner Parameter aus dem Modell getroffen werden sollen. Die Auswirkung einer hohen Multikollinearität ist, dass die Kostenwirkung der von Multikollinearität betroffenen Modellparameter nicht mehr präzise ausgedrückt werden kann. Dies hat zur Folge, dass die geschätzten Koeffizienten nicht alleine, sondern nur in Kombinationen aller Koeffizienten der kollinearen Modellparameter interpretiert werden können. Dabei ist zu berücksichtigen, dass große Standardfehler hervorgerufen durch Multikollinearitäten nicht ausschließlich durch einen Wegfall von kollinearen Parametern vermieden werden können. Ebenso kann eine Hinzunahme eines fälschlicherweise fehlenden Parameters die Varianz der Störterme reduzieren, was wiederum kleinere Varianzen aller Koeffizienten zur Folge hat (vgl. Kennedy, 2008, S. 200).

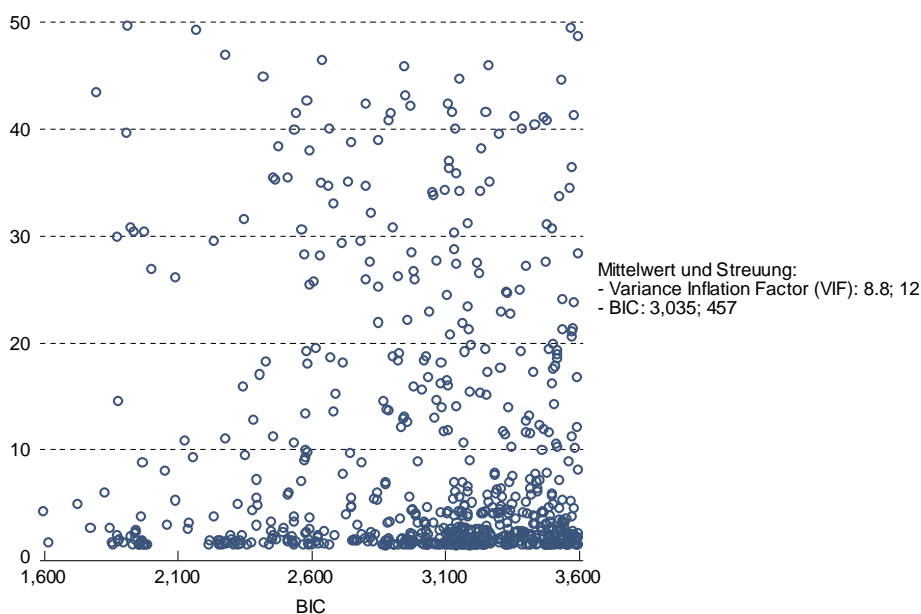
Andererseits hat Multikollinearität keine verzerrende Wirkung auf die Höhe der Koeffizienten und betrifft einzig die Parameter, welche kollinear zueinander sind. Diese Erkenntnis bedeutet, dass Multikollinearität auch keinen Einfluss auf das Bestimmtheitsmaß R^2 (und davon abgeleitete Gütemaße) und ebenso wenig auf die Höhe der Residuen der Schätzung hat (vgl. Kennedy, 2008, Kapitel 12.2). Genau diese Residuen, das heißt konkret die Kosten, welche durch die Strukturvariablen im Modell nicht erklärt werden können, sind jedoch Grundlage der Ermittlung eines Effizienzterms im Rahmen der Effizienzmessung. Eine empirische Untersuchung, welche nicht auf die kausalen Effekte einzelner Modellparameter auf die abhängige Variable fokussiert, sondern auf eine Vorhersage der abhängigen Variablen (hier: die Kosten) durch alle Strukturvariablen im Modell insgesamt, wird demnach nicht durch das Vorliegen von partieller Multikollinearität beeinträchtigt (vgl. Kennedy, 2008, S. 196).

Eine Auswahl von Vergleichsparametern anhand der statistischen Signifikanz steht im direkten Widerspruch zu der verstärkt definierten Anforderung aus der ARegV, die die Berücksichtigung der Heterogenität der Netzbetreiber weiter in den Vordergrund rückt.

Die zuvor dargestellten theoretischen Überlegungen, die durch entsprechende wissenschaftliche Literatur gestützt werden, lassen sich praktisch auch anhand des Datensatzes der BNetzA belegen: Es werden auf dieser Datengrundlage zufällig Modelle mit fünf bis neun Parametern

ausgewählt und damit in einem standardisierten Modell die Totex-Daten mittels OLS geschätzt. Nach Ausschluss statistisch unbrauchbarer Modelle wurde für insgesamt 710 Modelle jeweils der BIC-Wert und der maximale VIF (Variance Inflation Factor) berechnet. Abbildung 6 zeigt das Ergebnis, das besagt, dass kaum ein Zusammenhang zwischen BIC und VIF besteht: Hohe Multikollinearität führt nicht zu Modellen mit geringerem Erklärungsgehalt.

Parameter, die aus energiewirtschaftlicher Sicht im Modell berücksichtigt werden sollten, sind somit aufgrund möglicher Multikollinearität nicht vom Modell auszuschließen.



BMF-Graphik © Polynomics AG

Abbildung 6: Hohe Multikollinearität führt nicht zu schlechteren Modellen

Forderungen

Multikollinearität (solange sie nicht perfekt ist) verzerrt die Effizienzwertberechnung nicht. Daher ist diesem statistischen Kriterium sowie der alleinigen statistischen Signifikanz potentieller Vergleichsparameter bei der Beurteilung von Modellparametern kein starkes Gewicht zu geben.

Im Hinblick auf die Anforderungen der ARegV, Heterogenität in den Modellen abzubilden, müssen auch Parameterkombinationen, welche den gleichen oder ähnlichen Sachverhalt abbilden, zulässig sein.

Größere Modelle eignen sich besser, um die Heterogenitäten zwischen den Unternehmen adäquat abzubilden und sind daher unbedingt im Rahmen der Modellfindung zu berücksichtigen.

3.1.4 Problem der Pfadabhängigkeit und verzerrte Teststatistiken

Die BNetzA-Berater verfolgen bei der **Parameterauswahl drei Ansätze** (vgl. Folie 47):

- (1) Unrestringiertes Verfahren („Bottom-Up“ - BU)
- (2) Restringiertes Verfahren („Top-Down“ - TD) und
- (3) Weiterentwicklung Modell der zweiten Regulierungsperiode.

Das **BU-Verfahren** basiert auf einer schrittweisen Hinzunahme von Parametern und ist rein datenbasiert. Es werden zwei Stränge verfolgt, die Parameter werden in aggregierter und disaggregierter Form berücksichtigt. Das **TD-Verfahren** beruht auf einer aus ingenieurwissenschaftlicher Sicht vorpriorisierten Parameterliste. Aufbauend auf dieser Liste erfolgt ein sequentielles hierarchisches Vorgehen, in dem Parameterkandidaten einzeln geprüft werden. Bei dem **Ansatz 3** wird durch die Prüfung von bereits im Modell enthaltener und potentiell neuer Variablen eine Weiterentwicklung des Modells der zweiten Regulierungsperiode geprüft.

Laut den BNetzA-Beratern wird dieses Vorgehen gewählt, um das „Risiko bei konventioneller schrittweisen Regression zur Modellentwicklung“ einzugrenzen (Sumicsid et al. Folie 48). Es bleibt jedoch unklar, wie die Berater durch ihr Vorgehen dieses Risiko reduzieren. Auf Folie 49 wird aufgezeigt, dass angeblich die Methode „subset regression“ eingesetzt wird und für jede Modellgröße eine vollständige Suche mit Variablen um die Kombination mit tiefstem BIC-Wert erfolgt.

Laut Folie 77 (Sumicsid et al.) liegen dem BU-Ansatz bei der Verwendung von aggregierten Variablen 59 Parameter zur Verfügung und bei disaggregierten Variablen 426. Angenommen es erfolgt eine Beschränkung auf Modelle mit bis zu 10 Parametern, dann resultieren aus einem Set mit 59 möglichen Parametern näherungsweise 228 Milliarden mögliche Modelle. Durch die zusätzliche Berücksichtigung von disaggregierten Variablen steigt die Anzahl der möglichen Modelle exponentiell an. Dieser Fakt, sowie die (widersprüchliche) Darstellung des Vorgehens auf weiteren Folien, wie beispielsweise den Folien 49 und 51, legen die Vermutung nahe, dass im Rahmen der Parameterauswahl doch auf ein stufenweises Verfahren zur Parameterauswahl zurückgegriffen wurde. Damit bleibt unklar, wie die Berater das von ihnen selbst beschriebene Risiko der Pfadabhängigkeit durch ihr Vorgehen eingrenzen.

Der Einsatz von **stufenweisen Verfahren zur Modellfindung** wird in der wissenschaftlichen Literatur grundsätzlich sehr kritisch gesehen (Belloni et al., 2014, Copas und Long, 1991, Chatfield, 1995, Chatterjee und Hadi, 2015, Dersken und Keselman, 1992, Glantz und Slinker, 1990, Harrell, 2015, Keselman et al., 1998 und schließlich Kozbur, 2017) und führt nur unter sehr idealistischen Annahmen zur Identifikation des „wahren“ Modells (Wang, 2009).

Der Einsatz von stufenweisen Verfahren kann zu verzerrten Gütemaßen führen, dies erschwert die Gesamtbeurteilung eines Modells bzw. den Vergleich von verschiedenen Modellen. Gütemaße wie das korrigierte R², AIC (Akaike information criterion) oder BIC wurden ursprünglich dazu entwickelt, um eine eingeschränkte Anzahl an vorselezierten Modellen miteinander zu vergleichen (Grambsch und O'Brien, 1991, Harrell, 2015 und Keselman et al., 1998). Ähnliches gilt für die Verwendung von Teststatistiken wie beispielsweise F-Test und t-

Test. Beim Einsatz von stufenweisen Verfahren unterliegt die Teststatistik, die dem F-Test zu Grunde liegt, nicht der angenommenen Verteilung (Grambsch und O'Brien, 1991) und die geschätzten Standardfehler für den t-Test sind verzerrt (Altman und Andersen, 1989, und Kozbur, 2017). Hauptgrund für den Verlust an Aussagekraft der Gütemaße sowie der Teststatistiken ist, dass bei der Berechnung der Gütemaße oder Teststatistiken nicht korrigiert wird, dass bereits mehrere Modelle miteinander verglichen bzw. mehrere Parameter auf Signifikanz getestet wurden. Damit besteht die Gefahr, dass die resultierenden Regressionskoeffizienten verzerrt sind und die Auswahl von Parametern hauptsächlich vom geschätzten Koeffizienten und nicht vom „wahren“ Koeffizienten abhängig ist. Dies führt dazu, dass Koeffizienten, die überschätzt werden, mit einer höheren Wahrscheinlichkeit Eingang in das ausgewählte Modell finden (Chatfield, 1995, Dersken und Keselman, 1992, Greenland, 2000).

Sind die zur Auswahl stehenden Parameter zudem stark miteinander korreliert, wird die Identifikation des „wahren“ Modells zusätzlich erschwert. Eine hohe Korrelation zwischen Variablen führt dazu, dass die Variablen miteinander um die zu Grunde liegende Streuung konkurrieren (Glantz und Slinker, 1990, Copas und Long, 1991, Chatterjee und Hadi, 2015, Dersken und Keselman, 1992). Dadurch entstehen bei **stufenweisen Verfahren eine Pfadabhängigkeit** und damit eine gewisse Beliebigkeit bei der Auswahl der Modellparameter. Ob ein zusätzlicher Parameter in das Modell aufgenommen wird, ist stark abhängig davon, welche Parameter sich bereits im Modell befinden. Dieses Vorgehen führt in der Tendenz zu einer Wahl von Parametern, welche relativ unkorreliert miteinander sind. Heterogenitäten können dadurch nicht adäquat berücksichtigt werden. In der wissenschaftlichen Literatur wird das Verzerrungspotential eines fehlerhaft nicht aufgenommenen aber relevanten Vergleichsparameters wesentlich höher eingeschätzt als das Verzerrungspotential eines fehlerhaft aufgenommenen irrelevanten Vergleichsparameters (vgl. Smith, 1997, Galagedera, 2003, Kennedy, 2008, S. 93). Erschwerend kommt hinzu, dass fälschlicherweise ausgelassene relevante Parameter die Schätzkoeffizienten der weiteren erklärenden Variablen verzerren können (Kennedy, 2008, S. 101).

Vor dem Hintergrund der dargestellten **Schwächen von stufenweisen Verfahren** ist fraglich, ob eine Auswahl von Modellparametern auf hauptsächlichlicher Basis von statistischen Gütemaßen und Teststatistiken (Folie 49 BNetzA-Berater) zu einem sachgerechten Benchmarkingmodell führen kann.

Forderungen

Die Pfadabhängigkeit und die damit verbundenen verzerrten Gütemaße bei der Modellfindung nach dem Bottom-Up und dem Top-Down-Verfahren der BNetzA-Berater sind zu beseitigen.

Ein primäres Abstellen auf statistische Gütemaße und Teststatistiken ist aufgrund der Problematik der Pfadabhängigkeit nicht zielführend. Ingenieurwissenschaftliches Wissen sollte stärker berücksichtigt werden.

Als wesentlicher Baustein der Modellfindung ist auch das Modell der 2. Regulierungsperiode in Betracht zu ziehen.

Aufgrund der gestiegenen gesetzlichen Anforderungen zur Berücksichtigung von Heterogenitäten der Netzbetreiberaufgaben ist bei der Modellfindung eine Reduzierung der Parameteranzahl nicht sachgerecht.

3.1.5 Keine Berücksichtigung neuerer wissenschaftlich erprobter Ansätze

Das grundsätzliche Problem bei der **Modellauswahl** liegt darin, dass die Anzahl der möglicher Variablen weitaus höher ist als die Anzahl der Beobachtungen. Vor diesem Hintergrund beschränkt sich die BNetzA bei der **Variablenauswahl** auf eine einfache lineare Variablenspezifikation und schließt damit flexiblere funktionale Formen, wie beispielsweise quadratische Formen oder Interaktionen, a priori aus. Dies schränkt den Lösungsraum an optimalen Modellen stark ein und führt implizit zu der Annahme, dass das „wahre“ Modell linear ist.

Wie im Abschnitt 3.1.4. kritisch erläutert, setzt die BNetzA zur Modellfindung Verfahren wie „subset regression“ und „stepwise regression“ ein. Um die beschriebenen Schwächen wie Pfadabhängigkeit, multiples Testen, verzerrte Gütemaße oder verzerrte Koeffizienten zu mildern, wurden in der statistischen Theorie im Themengebiet „High-dimensional statistics“ Regularisierungsansätze (auch Shrinkage Verfahren genannt), wie beispielsweise „Lasso“ oder „Ridge“ entwickelt (Belloni et al., 2014 und Harrell, 2015, Bühlmann & Van De Geer 2011, Friedman et al. 2001). Diese Methoden wurden für das Szenarium entwickelt, dass die Anzahl der Variablen grösser als die Anzahl der Beobachtungen ist. Damit bieten diese Methoden grundsätzlich die Möglichkeit, die unterschiedliche funktionale Form von Modellen bereits bei der Parameterauswahl zu berücksichtigen und damit nicht implizit die Annahme zu treffen, dass das „wahre“ Modell linear ist.

Dieses Vorgehen ist grundsätzlich auch im Sinne der ARegV. Die Auswahl der Vergleichsparameter hat mit qualitativen, analytischen oder statistischen Methoden zu erfolgen, die dem Stand der Wissenschaft entsprechen (§ 13 ARegV). Daher sollten die BNetzA-Berater überprüfen, ob nicht noch weitere Methoden angewendet werden müssten, um eine Vorauswahl von Modellkandidaten vorzunehmen.

Forderung

Aufgrund der Schwächen der stufenweisen Verfahren zur Bestimmung von Effizienzmodellen sollten alternative Ansätze geprüft und deren Sensitivität ausgewiesen werden.

3.2 Wahl der Verteilung der Ineffizienz bei der SFA

Laut Auskunft der BNetzA-Berater im Rahmen des Workshops handelt es sich bei der **Wahl der Verteilung der Ineffizienz** um ein zweistufiges Vorgehen. Im ersten Schritt werden alle Modelle unter der Annahme einer „truncated normal distribution“ für die Verteilung der Ineffizienz geschätzt. Nach einer Vorauswahl an Modellen wird in einem zweiten Schritt auf Basis der auf einer unnummerierten Folie (vgl. Folie 63, Sumicsid et al.) beschriebenen Kriterien geprüft, welche Verteilung der Ineffizienz optimal ist (auch wenn auf dieser Folie von Kriterien zur Wahl der Verteilung des *stochastischen Störterms*, und nicht des Ineffizienzterms, die Rede ist).

Laut Kumbhakar & Lovell (2003, S. 90) hat die Wahl der Verteilung einen unumstrittenen Einfluss auf die Durchschnittseffizienz der Unternehmen in einem Sample. Folglich ist eine systematische Auswahl der Verteilung der Ineffizienz zu begrüßen. Grundsätzlich stehen **vier Verteilungen** zur Auswahl

- (1) „truncated normal distribution“
- (2) „half normal distribution“
- (3) „exponential distribution“
- (4) „gamma distribution“

Die vier Verteilungen unterscheiden sich durch ihre Anpassungsfähigkeit an die Daten und damit auch der Komplexität der Schätzung dieser Verteilungen. Nach Ritter & Simar (1997) werden mehrere tausend Beobachtungen benötigt, um komplexere Verteilungen der Ineffizienz wie beispielsweise die „gamma distribution“ oder die „truncated normal distribution“ robust zu schätzen. Liegen weniger Beobachtungen vor, wie es im Fall des Effizienzvergleichs der Verteilernetzbetreiber Strom mit ca. 200 Unternehmen der Fall ist, empfehlen Ritter & Simar (1997) sowie Kumbhakar & Lovell (2003) Verteilungen, wie beispielsweise die „exponential distribution“, da für diese Verteilung weniger Parameter zu schätzen sind als bei der „truncated normal distribution“. Wird dieser Umstand nicht berücksichtigt, besteht bei der Wahl einer komplexen Verteilung die Gefahr einer Scheingenauigkeit bei einer gleichzeitigen ungenauen Schätzung der Verteilung der Ineffizienz. Vor diesem Hintergrund ist das Vorgehen der BNetzA-Berater auf der ersten Stufe auf die „truncated normal distribution“ zurück zu greifen grundsätzlich zu hinterfragen. Vor dem Hintergrund der Erkenntnisse aus der wissenschaftlichen Literatur wäre die Auswahl der „exponential distribution“ in der ersten Stufe als solider einzuordnen. Bei der vorliegenden Datenlage ist zudem fraglich, in wie weit in einem zweiten Schritt überhaupt belastbar analysiert werden kann, ob es eine geeignetere Verteilung als die der „exponential distribution“ gibt.

Forderung

Bei der Annahme der Effizienzverteilung ist basierend auf der wissenschaftlichen Literatur bereits in der ersten Stufe auf eine Exponentialverteilung der Ineffizienz abzustellen.

3.3 Berücksichtigung unterschiedlicher Versorgungsaufgaben (Heterogenitäten)

Zur Sicherstellung der Abbildung der **Heterogenitäten von Aufgaben der Netzbetreiber** gemäß § 13 ARegV ist es notwendig, Vergleichsparameter auszuwählen, die den unterschiedlichen Aufgaben gerecht werden.

§ 13 Abs. 3 Satz 8 ARegV führt hierzu aus, dass „[d]urch die Auswahl der Vergleichsparameter (...) die strukturelle Vergleichbarkeit möglichst weitgehend gewährleistet sein und die Heterogenität der Aufgaben der Netzbetreiber möglichst weitgehend abgebildet [sollen].“

Unter Heterogenität von Aufgaben ist somit zu verstehen, dass strukturelle Gegebenheiten, die nicht für alle – möglicherweise auch nur für sehr wenige Netzbetreiber – vorliegen, zu signifikanten kostentreibenden Wirkungen führen.

Durch das Abstellen in der von den BNetzA-Berater durchgeführten Kostentreiberanalyse auf vorwiegend Kriterien (insbesondere dem BIC), die die Anzahl der Parameter systematisch minimieren, können wesentliche Kostentreiber für bestimmte Gegebenheiten bzw. Versorgungsaufgaben, von denen jeweils nur wenige Netzbetreiber betroffen sind, nicht hinreichend identifiziert werden. Weiter darf auch durch das im Workshop geäußerte Bestreben der BNetzA-Berater, eine möglichst geringe Anzahl von Vergleichsparametern im finalen Effizienzvergleichsmodell zu verwenden, nicht dazu führen, dass insbesondere aggregierte Parameter verwendet werden und somit das Risiko besteht, dass die Heterogenität der Netzbetreiberaufgaben beim Effizienzvergleich nicht sachgerecht erfasst wird.

Denn beispielsweise wirkt sich die Heterogenität bei der Anzahl der vom Netzbetreiber betriebenen Spannungsebenen auf den Effizienzvergleich aus. Grund: dem regulatorischen Konzept folgend ist es maßgebend, welche Kosten in den Effizienzvergleich einzubeziehen sind. Hat ein Netzbetreiber eigene Hoch- oder Mittelspannungsnetze, dann sind diese Kosten Teil des Effizienzvergleichs. Hat ein Netzbetreiber diese Netzebenen nicht, entstehen zwar auch Kosten in Form von vorgelagerten Netzkosten. Diese vorgelagerten Netzkosten sind aber für den Benchmark nicht relevant, da diese als dauerhaft nicht beeinflussbar klassifiziert werden und von der weiteren Betrachtung ausgeschlossen sind. Wenn diese Unterschiede im Effizienzvergleich nicht adäquat aufgrund von aggregierten oder einer minimierten Anzahl von Parametern berücksichtigt werden, kann die Folge sein, dass insbesondere kleinere Netzbetreiber, die nur in der Niederspannung tätig sind, den Benchmark für größere Netzbetreiber setzen, die aber benchmarkrelevant alle Spannungsebenen bedienen. Dieses Problem kann nicht durch die Aggregation von Parametern gelöst werden, da sich die Kosten in den unterschiedlichen Spannungsebenen unterscheiden (z.B. Kosten pro Kilometer). Die bislang vorgeschlagenen Modelle „BU1+“ und „TD2+“ – mit einer geringen Anzahl von Parametern – bilden eben diese Unterschiede in der Versorgungsaufgabe nicht ab. In diesen Modellen wird beispielsweise die aggregierte Leitungslänge über alle Netzebenen angesetzt, was zu erheblichen **Verzerrungen der Effizienzwerte** führen kann.

Diese **Heterogenität der Aufgaben von Netzbetreibern wird auch durch die Analyse der von der BNetzA zur Verfügung gestellten Daten deutlich** und führt bei einer Verwendung von aggregierten Parametern dazu, dass für die am Effizienzvergleich teilnehmenden Netzbetreiber die – historisch begründeten – teilweise deutlichen strukturellen Unterschiede und Versorgungsaufgaben nicht sachgerecht erfasst werden.

Des Weiteren ist mit Blick auf die explizit vorgeschlagenen Parameter „yinstalledPower.NE4to7.sum“ bzw. „yinstalledPower.renewables.tot“ der Modelle „BU1+“ bzw. „TD2+“ fraglich, ob die vorgeschlagenen Parametrisierung eine angemessene Ausgewogenheit mit sich bringt, um die Heterogenität zwischen den unterschiedlich vom Ausbau der dezentralen Erzeugung betroffenen Verteilernetzbetreibern sachgerecht abzubilden. Denn bereits im Jahr 2014 hatte der BNetzA-Berater (IAEW) in seinem Forschungsprojekt „Moderne Verteilernetze für Deutschland“⁴ (BMW i Verteilernetzstudie) deskriptiv dargelegt, dass „[r]und 75 % aller Netzbetreiber (...) bereits heute zumindest punktuell vom Netzausbau durch die Integration

⁴ BMW i-Studie „Moderne Verteilernetze für Deutschland“; E-Bridge, IAEW, Offis; 12. September 2014; https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/verteilernetzstudie.pdf?__blob=publicationFile&v=5

von EE-Anlagen betroffen [sind], wenige ländliche Netzbetreiber sind massiv und in der Fläche betroffen.“ Weiter führt der Berater in der Studie aus, dass „der resultierende Netzausbaubedarf (...) sich besonders in der Niederspannungsebene [unterscheidet]. Hier ist der Netzausbau wesentlich durch die angeschlossene PV-Leistung getrieben.“, um schließlich festzustellen: „Die höchste Anzahl an EE-Anlagen ist dabei an die Niederspannungsnetze angeschlossen. Ca. 1.000.000 Photovoltaikanlagen sind an Niederspannungsnetze angeschlossen und haben eine installierte Erzeugungsleistung von insgesamt ca. 16 GW. (...) In der Mittelspannungsebene ist die höchste Einspeiseleistung aus Erneuerbaren Energien [14,5GW Wind und 5,5GW Photovoltaik] angeschlossen.“

Vor diesem Hintergrund ist fachlich nicht ausreichend ausgeführt, wieso beispielsweise das Modell „BU1+“ den Fokus auf die Spannungsebenen 4 bis 7 legt und somit Verteilernetzbetreiber mit Netzanschlüssen von erneuerbaren Energien mit nahezu vergleichbaren installierten Leistungen in den unteren Spannungsebenen unberücksichtigt lässt. Denn die Studie belegt, dass der Zuwachs von dezentraler Erzeugung auch kurzfristig erhebliche Mehrkosten in den Stromverteilernetzen verursacht. Allerdings geht die Studie davon aus, dass diese Kostenwirkung aufgrund der heterogenen Struktur der Verteilernetze und des lokal differierenden Zubaus von EE-Anlagen je nach Regionen und Spannungsebenen höchst unterschiedlich ausfällt.

Die BNetzA-Berater (IAEW) selbst waren der Ansicht, dass der **Effizienzvergleich dieser zunehmenden Heterogenität der Verteilernetzbetreiber gerecht werden** muss. Die selektive Abbildung einer einzigen Erzeugungsart (Photovoltaik) in einer einzigen Spannungsebene (Mittelspannung) wie im Modell „RP2+“ bzw. die aggregierte installierte Erzeugungsleistung aller Erzeugungsanlagen („BU1+“) oder die installierte Erzeugungsleistung aller EEG-Erzeugungsanlagen („TD2+“) bildet diese Heterogenität somit noch nicht ab.

Des Weiteren ist zu hinterfragen, wieso das Modell „TD2+“ unterstellt, dass sich die Netzkosten aller Netzbetreiber auch durch die Anschlussleistungen von erneuerbaren Energien erklären lassen, wenn bspw. laut BMWi-Verteilernetzstudie 276 Verteilernetzbetreiber vom Ausbau der erneuerbaren Energien jedenfalls in der Niederspannung nahezu nicht betroffen sind (vgl. S.8 der BMWi Verteilernetzstudie). Hier zeigt sich also Heterogenität in mehrfacher Hinsicht: die Netzbetreiber sind insgesamt unterschiedlich vom Ausbau der EEG-Anlagen betroffen und selbst innerhalb der betroffenen Netzbetreiber gibt es relevante Unterschiede bezüglich der betroffenen Spannungsebene, Energiedargebot und Verhältnis der Einspeisung zur Last.

Insofern ist das Heranziehen dezentraler, erzeugungsorientierter Parameter zur Bestimmung der spezifischen Auswirkungen der Energiewende auf die Netzkosten grundsätzlich zielführend. Jedoch ist dieses Vorgehen zur Berücksichtigung der Heterogenität zwischen Netzbetreibern mit verschieden ausgeprägten Versorgungsaufgaben noch nicht hinreichend, da die Energiewende und die bestehenden als auch zukünftigen Herausforderungen im Netzbetrieb bei den im Effizienzvergleich vertretenen Netzbetreibern unterschiedliche Folgen haben.

Weiter ist im Zusammenhang mit der Heterogenität anzuführen, dass bei einer ersten Beurteilung der von den BNetzA-Beratern präsentierten Grundmodelle sich einige Fragen hinsichtlich deren **energiewirtschaftlichen Plausibilität** stellt. So findet sich beispielsweise der Parame-

ter Höchstlast nicht im Modell „BU1+“, obwohl dieser von den BNetzA-Beratern „als aus ingenieurwissenschaftlicher Sicht besonders gut“ beurteilt wurde, „um die Kosten der Netzbetreiber abzubilden“ (Sumicsid et al. Folie 26). Des Weiteren ist im Gegensatz zum Modell der 2. Regulierungsperiode die Netzlänge auf der Mittelspannungsebene nicht differenziert nach Kabeln und Freileitungen ausgewiesen, obwohl diese beiden Verlegungsvarianten mit unterschiedlichen Kosten verbunden sind.

Die zwischen den Verteilernetzbetreiber im Datensatz der BNetzA zu beobachtende Heterogenität kann auch nicht über die Ausnahmenvorschrift des § 15 Abs. 1 ARegV gelöst werden. § 13 steht systematisch in der ARegV vor § 15, d.h. die Belange und Aufgaben des § 13 sind durch die Behörde und ihre Berater zu lösen; ein Nichtbetrachten der Heterogenität unter implizitem Verweis, auf die Möglichkeit, einen Härtefallantrag zu stellen, genügt dem gesetzlichen Auftrag nicht. Diese lediglich in Einzelfällen zur Anwendung gelangende Vorschrift bezieht sich – insbesondere auch wegen der einschränkenden Auslegung durch die BNetzA – auf individuell bestehende Besonderheiten der Versorgungsaufgabe und nicht etwa auf strukturelle Unterschiede der Verteilernetzbetreiber.

Zusammenfassend ist daher festzuhalten, dass **eine ausreichende Anzahl geeigneter Parameter bereits in den Basismodellen notwendig ist**, um sicherzustellen, dass die Heterogenität bei den Aufgaben aller Netzbetreiber gemäß § 13 ARegV angemessen einbezogen werden kann (vgl. Kap:1).

Forderung

Den im Datensatz der BNetzA dargestellten strukturellen Unterschieden zwischen den Netzbetreibern ist bei der Anzahl und Wahl der Parameter (aggregiert und disaggregiert) adäquat Rechnung zu tragen.

3.4 Unterschiede DEA und SFA/Parameterwahl

Wie in den bisherigen Effizienzvergleichen geht aus den Konsultationsunterlagen hervor, dass die Berater auch für den dritten Effizienzvergleich der Stromverteilernetzbetreiber die Verwendung der **identische Parameter für SFA und DEA** planen. Dies entspricht einer selbstauferlegten Restriktion und wird von der ARegV nicht gefordert. Es besteht allerdings die Gefahr, dass durch diese Restriktion die schlechtesten Eigenschaften beider Methoden kombiniert werden und die Möglichkeiten bzw. das Potential der methodenimmanenten Vorteile ungenutzt bleibt. Insofern birgt dieser selbstauferlegte Zwang zur Verwendung der exakt selben Parameter auch die Gefahr, die gesetzliche Vorgabe des „best-of-four-Benchmarks“ auszuhöhlen, denn offensichtlich dominiert die (für die Kostentreiberanalyse entscheidendere) SFA die DEA zunehmend. Vor allem zwei Aspekte stehen hier im Vordergrund:

1. Verzerrungspotential bei der DEA bei fälschlicherweise nicht berücksichtigten Parametern

Die **Methode der DEA** selbst beinhaltet in ihrer ursprünglichen Form, und im Gegensatz zu Regressionsanalysen wie OLS bzw. SFA, keine Leitlinien zur Auswahl von relevanten Ver-

gleichparametern bzw. Kostentreibern. Teilweise wird sogar die Anwendung von Regressionsanalysen zur Identifikation von Vergleichsparametern in der DEA wird als kritisch angesehen (Ruggiero, 1998, S. 370). Die wissenschaftliche Literatur zeigt weiter, dass in der DEA ein Vergleichsparameter, dessen Relevanz unsicher ist, im Zweifelsfall eher im Modell berücksichtigt als ausgeschlossen werden sollte (vgl. Smith, 1997, S. 250). Das Verzerrungspotential eines möglicherweise zu großen Modells ist als kleiner einzuschätzen als dasjenige einer Unterspezifikation (Vgl. Galagedera und Silvapulle, 2003, S. 657). Dieses Risiko von statistischen Fehlern erster und zweiter Ordnung sollte gegeneinander aufgewogen werden. Im Zweifelsfall (zum Beispiel, wenn ein Parameter in der Kostenprüfung eine sehr hohe Kostenwirkung hat oder in einer Vielzahl von Schätzungen relevant erscheint) ist ein Parameter für die DEA eher aufzunehmen, als zu verwerfen – auch wenn eine Verwendung in das **SFA-Modell** zum Beispiel aufgrund fehlender Konvergenz kritisch erscheint. Ein zusätzliches Verzerrungspotential existiert, wenn konstante Skalenerträge unterstellt werden (vgl. Galagedera und Silvapulle, 2003, S. 657): Da diese in der ARegV vorgegeben werden, ist es umso wichtiger, dass das Risiko von fehlerhaft nicht berücksichtigten Parametern minimiert wird.

Aufgrund der Eingeschränktheit der Möglichkeiten der Identifikation von Vergleichsparametern in der DEA und der erwähnten Verzerrungspotentiale durch fehlende aber relevante Vergleichsparameter werden in neueren wissenschaftlichen Untersuchungen Methoden zur Identifikation von Strukturvariablen innerhalb der DEA beschrieben. Beispielsweise können mittels sogenannten Bootstrap-Verfahren (vgl. z. B. Simar und Wilson, 2001) die Relevanz von Vergleichsparametern in einem DEA-Modell geprüft werden. Dasselbe Verfahren kann auch angewandt werden, um relevante Vergleichsparameter zusammenzufassen bzw. gewichtet zu aggregieren, damit die Heterogenität der Netzbetreiber adäquat berücksichtigt werden kann. Neuere Untersuchungen empfehlen auch spezifische Vorgehensweisen bei hoch korrelierten Variablen und kleineren Stichproben, wie beispielsweise das sogenannte Efficiency Contribution Measure (Nataraja und Johnson, 2011, für eine Übersicht).

2. SFA und DEA unterscheiden sich in Bezug auf die Modellierung des Zusammenhangs zwischen den berücksichtigten Parametern und den Kosten

Durch die Novellierung der ARegV im Sommer 2016 ist die BNetzA verpflichtet, die Heterogenität der Aufgaben der Netzbetreiber sachgerecht im Effizienzvergleichsverfahren abzubilden. In der SFA kann dies unter anderem (aber nicht nur) durch die Wahl einer flexibleren funktionalen Form gewährleistet werden, wie dies im Rahmen des Effizienzvergleichsverfahrens für Gasverteilernetzbetreiber zur dritten Regulierungsperiode mit Anwendung der Translog-Funktion geschehen ist. Die **Translog-Funktion** hat dabei den Vorteil, dass diese Spezifizierung eher geeignet ist, das Problem von **Heteroskedastizität** abzufangen. Die Flexibilisierung der Funktion ist deshalb wichtig, weil bei der **SFA** genau wie bei der OLS für alle Unternehmen die gleiche Kostenwirkung pro Parameter unterstellt wird. Dies im Gegensatz zur DEA, bei der für jedes Unternehmen seine eigene Wirkung eines Parameters auf die Kosten berücksichtigt wird.

Aufgrund der **nichtparametrischen Eigenschaft der DEA** unterliegt diese Methode keinen Restriktionen bezüglich funktionaler Form. Die Berater der BNetzA würdigen die methodenimmanenten Unterschiede zwischen SFA und DEA jedoch nicht, indem der Korrelation der Effizienzwert aus DEA und SFA ein hohes Gewicht bei der Modellauswahl zugesprochen wird (vgl. Sumicsid et al. Folie 87) und die Parameterwahl zwischen DEA und SFA identisch sein muss. Dadurch werden an mindestens zwei Stellen im Prozess die Einschränkungen der SFA auf die DEA übertragen: Erstens können nur Modellkandidaten gewählt werden, welche in der SFA konvergieren, d. h. sollte es aufgrund der Anwendung einer flexibleren funktionalen Form in der SFA notwendig erscheinen, aufgrund von Konvergenzproblemen weniger Parameter aufzunehmen, kann dies negative Auswirkungen auf Effizienzermittlung mittels DEA haben. Die Parameterauswahl für ein konvergierendes SFA-Modell darf in der Folge nicht zur Anwendung eines unterspezifizierten DEA-Modells führen – denn die Abbildung von Heterogenität ist in beiden Berechnungsmethoden erforderlich. Zweitens führt die selbstauferlegte Restriktion auf linear Modelle mit Niveauvariablen im Prozess der Kostentreiberanalyse zu einer Beschneidung der Vorteile der DEA, in welcher keine Parametrierungen vorgenommen werden müssen.

Forderungen

Gemäß den Vorgaben der ARegV handelt es sich sowohl bei der SFA als auch der DEA um gleichberechtigte Berechnungsmethoden, deren Stärken zur Durchführung eines sachgerechten Effizienzvergleichsverfahrens zwingend genutzt werden sollten. Falls dieses im Ergebnis zu einer unterschiedlichen Parametrierung der Methoden bzw. zu separaten methodenspezifischen und methodenangepassten Kostentreiberanalysen führen sollte, wäre es im Einklang mit den gesetzlichen Vorgaben.

Die selbstauferlegte Restriktion, die gleichen Vergleichsparameter für die DEA und SFA zu verwenden und der Vergleichbarkeit der Effizienzwert aus den beiden Methoden ein hohes Gewicht gegeben wird, konterkariert die in der ARegV vorgesehene Bestabrechnung über beide Methoden.

3.5 Durchführung der Ausreißeranalyse

Identifikation von Ausreißern in der SFA

Gemäß Aussagen der Berater der BNetzA werden **Ausreißer** in parametrischen Methoden mittels des Kriteriums der Cook's Distance identifiziert (vgl. Sumicsid et al. Folie 56). Weitere Ausführungen zur Ausreißerproblematik wurden an der Konsultation nicht getätigt. In diesem Zusammenhang ist es wichtig, dass durch die Auswahl des entsprechenden Grenzwertes der zugenommenen Heterogenität von Aufgaben der Netzbetreiber Rechnung getragen wird. Aufgrund der Wichtigkeit der Ausreißeranalysen und der Nichtexistenz eines einzigen korrekten Grenzwertes in der wissenschaftlichen Literatur sollten Sensitivitätsanalysen sowohl bezüglich der Grenzwerte als auch der weiteren in der ARegV genannten Methoden zur Identifikation von Ausreißern in parametrischen Methoden (insb. DFBETAS und DFFITS) durchgeführt und ausgewiesen werden.

Außerdem muss festgehalten werden, dass für die Durchführung eines Effizienzvergleichs eine **vergleichbare Datenbasis von objektiv strukturell vergleichbaren Unternehmen** sowohl gesetzlich gefordert (§ 21a Abs. 2 Satz 4 EnWG) als auch statistisch unabdingbar ist. Die BNetzA selbst hat der im Prozess der Evaluierung und Weiterentwicklung der Anreizregulierung vorgeschlagen, in Zukunft Unternehmen mit extremen Werten der **Cook's Distance** sowohl für die SFA als auch für die DEA als Ausreißer zu klassifizieren und auszuschließen. Dies muss als ein dem eigentlichen Effizienzvergleich vorgelagerter Schritt verstanden werden. Erst nach diesem Ausschluss der von strukturell nicht vergleichbaren Unternehmen, kann die **reguläre Ausreißerkorrektur**, wie sie in der ARegV angelegt ist, erfolgen.

Es ist im Weiteren genau zu dokumentieren, welchen Cook's Distance-Wert auffällige Unternehmen zugewiesen bekommen und außerdem, wie die Cook's Distance auffälligen Unternehmen ihren Effizienzwert erhalten.

Überdies hat sich in der **SFA** gezeigt, dass einzelne Unternehmen die Lageparameter der geschätzten Verteilung für Ineffizienz stark beeinflussen. Dies hat einen starken Einfluss auf die Aufteilung der Residuen in „Rauschen“ und „Ineffizienz“. Neben einer Analyse der Cook's Distance Werte (d. h. des Einflusses auf den deterministischen Teil der Schätzgerade) ist ebenfalls der Einfluss einzelner Unternehmen auf die Verteilparameter zu untersuchen. Nur dieses Vorgehen wäre ARegV-konform und stellt zudem eine weitere Annäherung an den Stand von Wissenschaft dar.

Identifikation von Ausreißern in der DEA

Gemäß den Berichten zu den bisher durchgeführten Effizienzvergleichen der Stromverteilernetzbetreiber wurde für die **Identifikation von Ausreißern** in der DEA im Rahmen der Dominanzanalyse ein F-Test durchgeführt. Aus der Konsultationspräsentation (vgl. Sumicsid et al. Folie 55) wird nicht klar, ob auch bei diesem Effizienzvergleich dieser Test angewendet wird. Dieser Test ist aus zwei Gründen im vorliegenden Fall jedoch grundsätzlich nicht anwendbar: Zum einen stammen die Effizienzwerte aus einer nicht-parametrischen Untersuchung und es ist daher verwunderlich, dass ein parametrischer Test angewendet wird, welcher voraussetzt, dass die Effizienzwerte aus einer gestutzten Normalverteilung stammen. Zum anderen geht der Test von der Annahme aus, dass die beiden zu vergleichenden Werte aus zwei unabhängigen Stichproben stammen (vgl. Banker, 1996). Diese Annahme ist im vorliegenden Fall klar verletzt: Es werden zwei verschiedene Effizienzwerte des gleichen Unternehmens miteinander verglichen. Die Dominanzanalyse sollte daher auf Basis nicht-parametrische Tests durchgeführt werden, welche die „paired“-Struktur der vorliegenden Daten berücksichtigt (z. B. Wilcoxon-Test bzw. Mann-Whitney-Test, Sign-Tests, Bootstrapping-Methoden o. ä.).

Die Tatsache, dass die bisher von den BNetzA-Beratern angewandte Dominanzanalyse zu Verzerrungen in der Ermittlung der DEA-Effizienzwerte resultieren kann, wird beispielhaft in Abbildung 7 dargestellt. Anhand des von den BNetzA-Beratern vorgestellten Grundmodells „RP2+“ und basierend auf Kosten nach § 14 STromNEV werden auf der Y-Achse einerseits die DEA-Effizienzwerte ermittelt mit einem F-Test in der Dominanzanalyse, auf der X-Achse andererseits diejenigen ermittelt mit einem Mann-Whitney-Test in der Dominanzanalyse darge-

stellt. Dabei werden im ersten Fall insgesamt 11 Ausreißer (aus Dominanz- und Supereffizienzanalyse), im zweiten Fall insgesamt 18 Ausreißer identifiziert und ausgeschlossen. Durch die Identifikation von 7 zusätzlichen Ausreißern steigen die DEA-Effizienzwerte signifikant an.

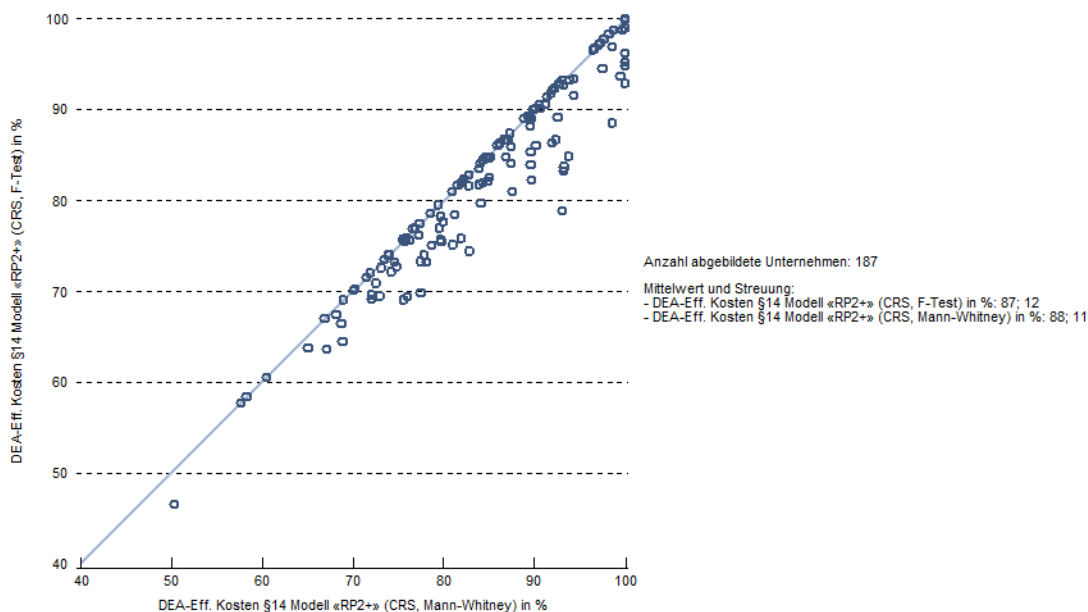


Abbildung 7: DEA-Effizienzwerte mit F-Test und Mann-Whitney-Test in der Dominanzanalyse

Das bisherige Verfahren zur Supereffizienzanalyse zeigt deutliche Mängel aufgrund der Problematik von „maskierten“ Ausreißern. Der BGH bestätigt die Notwendigkeit einer Identifizierung von Netzbetreibern, die in ihrer Versorgungsaufgabe gerade nicht vergleichbar sind.

Wird auf den von der BNetzA publizierten Daten und mit dem Modell der „RP2+“ eine Supereffizienzanalyse durchgeführt, wird ein Unternehmen mit einem Supereffizienzwert von 300% identifiziert. Wird dieses strukturell offensichtlich spezielle Unternehmen vorgängig ausgeschlossen, erhalten andere Unternehmen hohe Supereffizienzwerte von bis zu 250% (unterhalb der Winkelhalbierenden in der Abbildung 8, in welcher die „Übereffizienzwerte“, d.h. die DEA-Effizienzwerte über 100% dargestellt sind). Zusätzlich werden neue Unternehmen als supereffizient erkannt und somit als mögliche Ausreißer identifiziert mit der möglichen Konsequenz einer Verzerrung der Effizienzwerte.

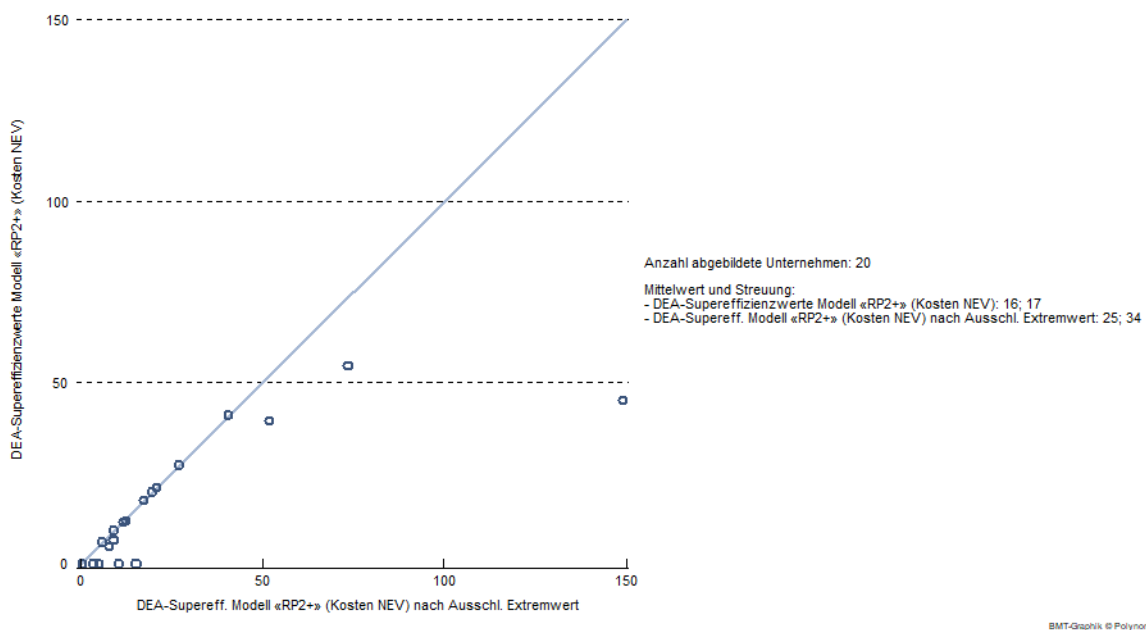


Abbildung 8: Ausreißeranalyse und Problem von „maskierten“ Ausreißern

Dieses Vorgehen ist nach Ansicht der Verbände aus den Vorgaben der Verordnung ableitbar. Es kommt auch nicht zu einem endlosen „Abschälen der Zwiebel“, denn die Supereffizienzwerte aller Unternehmen normalisieren sich, wenn sämtliche Ausreißer aus der Berechnung entfernt werden.

Forderungen

Das Problem verdeckter Ausreißer („masking effects“ und „swamping effects“) ist in der Literatur bekannt. Die bisherigen Ausreißeranalysen sind ungeeignet, um multiple Ausreißer zu entdecken.

Die bei der Ausreißeranalyse verwendeten Teststatistiken sollten die notwendigen Annahmen erfüllen. Insbesondere die Teststatistiken zur Dominanzanalyse erfüllen diese derzeit nicht.

Eine dem Stand der Wissenschaft und der aktuellen Rechtsprechung entsprechende Ausreißeranalyse ist daher zwingend erforderlich.

3.6 Second-Stage-Analysen/Sensitivität

In der Vergangenheit wurden die Ergebnisse mit Hilfe von **Second-Stage-Analysen** plausibilisiert. Auch für diesen Effizienzvergleich ist dieses Vorgehen grundsätzlich vorgesehen. Aus der wissenschaftlichen Literatur geht hervor, dass die Second-Stage-Analyse nicht zur Modellvalidierung oder –plausibilisierung entwickelt wurde, in der DEA – wenn überhaupt – nur für Umweltparameter zulässig ist (vgl. Banker, 2008, Simar, 2007 und 2011, sowie Johnson, 2012) und in der SFA grundsätzlich nicht angewendet werden sollte (vgl. Wang, 2002).

Die meisten Unternehmen erhielten in der Vergangenheit ihren bestabgerechneten Effizienzwert aus der SFA. Die Aufteilung der Restgröße in Ineffizienz und Störterm basiert auf den geschätzten Varianzen σ_u und σ_z . Diese Faktoren beeinflussen die Effizienzwerte aller Unternehmen in einem ähnlichen Ausmaß. Fehler bei der Modellfindung, insbesondere dann, wenn sie diese Schätzparameter stark beeinflussen, können durch eine Second-Stage-Analyse grundsätzlich nicht aufgedeckt werden. Es ist daher anzuraten, insbesondere für Modellentscheidungen, welche nicht durch eindeutige statistische Tests untermauert werden können (z. B. Abbildung Heteroskedastizität), Sensitivitätsanalysen durchzuführen. Hierbei werden die Effizienzwerte unterschiedlicher Modellvarianten miteinander verglichen (z. B. durch Scatterplots oder geeignete Testverfahren). Zudem können so systematische Einflüsse sichtbar gemacht werden, was Second-Stage-Analysen nicht leisten können. Mit Hilfe dieser Sensitivitätsanalysen wird man zudem wesentlich besser den Anforderungen aus § 21a Abs. 5 EnWG gerecht: „Die Methode zur Ermittlung von Effizienzvorgaben muss so gestaltet sein, dass eine geringfügige Änderung einzelner Parameter der zugrunde gelegten Methode nicht zu einer, insbesondere im Vergleich zur Bedeutung, überproportionalen Änderung der Vorgaben führt.“

Forderung

Die Second-Stage-Analysen dürfen gemäß wissenschaftlicher Erkenntnis weder für die SFA-Methode noch zur Modellvalidierung respektive -plausibilisierung in der DEA eingesetzt werden. Geeigneter sind Sensitivitätsanalysen, welche die Ergebnisse verschiedener Effizienzmodellrechnungen miteinander vergleichen

Literatur

- Altman, D. und P. K. Andersen, 1989, Bootstrap investigation of the stability of a Cox regression model, *Statistics in medicine* 8.7, 771-783.
- Banker, R. und R. Natarajan, 2008, Evaluating contextual variables affecting productivity using data envelopment analysis, *Operations Research*, 56(1), 48-58.
- Banker, R., 1996, Hypothesis tests using data envelopment analysis, *Journal of Productivity Analysis*, 7, 139-159.
- Belloni, A., V. Chernozhukov und C. Hansen, 2014, Inference on Treatment Effects after Selection among High-Dimensional Controls, *Review of Economic Studies*, Vol. 81, No. 2, 608-650.
- Bühlmann, Peter, and Sara Van De Geer. *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media, 2011.
- Chatfield, C., 1995, Model Uncertainty, data mining and statistical inference (with discussion), *Journal of the Royal Statistical Society, series B* 158, 419-466.
- Chatterjee, S. und A. S. Hadi, 2015, *Regression analysis by example*, John Wiley & Sons.
- Copas, J. B. und T. Long, 1991, Estimating the residual variance in orthogonal regression with variable selection, *The Statistician*, 51-59.
- Derksen, S., und H. J. Keselman, 1992, Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables, *British Journal of Mathematical and Statistical Psychology* 45.2, 265-282.
- Dyson, R.G., R. Allen, A.S. Camanho, V.V. Podinovski, C.S. Sarrico und E.A. Shale, 2001, Pitfalls and protocols in DEA, *European Journal of Operational Research* 132, 245–259.
- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*. Vol. 1. No. 10. New York, NY, USA: Springer series in statistics, 2001.
- Galagedera, D. und P. Silvapulle, 2003, Experimental evidence on robustness of data envelopment analysis, *Journal of the Operational Research Society*, 54, 654-660.
- Glantz, S. A. et al., 1990, *Primer of applied regression and analysis of variance*, McGraw Hill Book Co.
- Grambsch, P. M., und P. C. O'Brien, 1991, The effects of transformations and preliminary tests for non-linearity in regression, *Statistics in Medicine* 10.5, 697-709.
- Greenland, S., 2000, When should epidemiologic regressions use random coefficients?, *Biometrics* 56.3, 915-921.
- Harrell, F., 2015, *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*, Springer.
- Johnson, A. L. und T. Kuosmanen, 2012, One-stage and two-stage DEA estimation of the effects of contextual variables, *European Journal of Operational Research*, 220(2), 559-570.
- Kennedy, P., 2008, *A Guide to Econometrics* 6. ed., Wiley-Blackwell.

- Keselman, H. J., et al., 1998, A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements, *Communications in Statistics-Simulation and computation* 27.3, 591-604.
- Kozbur, D., 2017, Testing-Based Forward Model Selection, *American Economic Review* 107.5, 266-69.
- Kumbhakar, S. und C. Lovell, 2000, *Stochastic Frontier Analysis*, Cambridge: Cambridge University Press.
- Kumbhakar, Subal C., and CA Knox Lovell. *Stochastic frontier analysis*. Cambridge university press, 2003.
- Nataraja, N., und A. L. Johnson, 2011, Guidelines for Using Variable Selection Techniques in Data Envelopment Analysis, *European Journal of Operational Research*, December 2011.
- Ritter, Christian, and Léopold Simar. "Pitfalls of normal-gamma stochastic frontier models." *Journal of productivity analysis* 8.2 (1997): 167-182.
- Ruggiero, J., 1998, A new Approach for Technical Efficiency Estimation in Multiple Output Production, *European Journal of Operational Research* 111, 369–380.
- Simar, L. und P. W. Wilson, 2007, Estimation and inference in two-stage, semiparametric models of production processes, *Journal of Econometrics*, 136(1), 31-64.
- Simar, L. und P. W. Wilson, 2011, Two-stage DEA: caveat emptor, *Journal of Productivity Analysis*, 36.
- Simar, L., und P. W. Wilson, 2001, Testing Restrictions in Nonparametric Efficiency Models, *Communications in Statistics - Simulation and Computation* 30,159–184.
- Smith, P., 1997, Model Misspecification in Data Envelopment Analysis, *Annals of Operations Research*, 73(1), 233-252.
- Smith, P., 1997, Model Misspecification in Data Envelopment Analysis, *Annals of Operations Research* 73(1), 233–252.
- Sumicsid et al., 2018, Konsultation zur Auswahl der Vergleichsparameter zum Effizienzvergleich der deutschen Elektrizitätsverteilnetzbetreiber, Workshop vom 25.07.2018, Bonn.
- Wang, H. J. und P. Schmidt, 2002, One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels, *Journal of Productivity Analysis*, 18(2), 129-144.
- Wang, H., 2009, Forward regression for ultra-high dimensional variable screening, *Journal of the American Statistical Association* 104.488, 1512-1524.
- White, H. L., Jr., 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for