

Stellungnahme der Netze BW GmbH zur Branchenkonsultation zum Effizienzvergleich Gas für die vierte Regulierungsperiode

Stuttgart 27. Januar 2023

Ein Unternehmen der EnBW



1 Einleitung

Das Referat „Anreizregulierung und Vergleichsverfahren“ der Bundesnetzagentur hat am 16. Dezember 2022 eine Konsultationsveranstaltung zum Effizienzvergleich der deutschen Gasverteilernetzbetreiber für die vierte Regulierungsperiode durchgeführt und die Netzbetreiber gebeten, bis zum 27. Januar 2023 Stellung zu nehmen.

Die ARegV sieht vor, dass die Bundesnetzagentur die betroffenen Wirtschaftskreise und Verbraucher zu den Methoden des Effizienzvergleichs und zur Auswahl der Vergleichsparameter anhört (§ 12 Abs. 1 S. 2 und § 13 Abs. 3 S. 10 ARegV).

Gerne nimmt die Netze BW die Möglichkeit zur Stellungnahme wahr. Grundlage für unsere Stellungnahme sind die Ausführungen während der digital durchgeführten Konsultationsveranstaltung sowie die dort und im Nachgang bereitgestellten Präsentationen.

Gesetzlich ist das Benchmarkingverfahren in §21a Satz 5 EnWG verankert: So müssen die Effizienzvorgaben so gestaltet und über die Regulierungsperiode verteilt sein, dass sie erreicht oder übertroffen werden können. Dabei dürfen geringfügige Änderungen einzelner Parameter der zugrunde gelegten Methode nicht zu einer, insbesondere im Vergleich zur Bedeutung, überproportionalen Änderung der Vorgaben führen. Objektive strukturelle Unterschiede sowie die bestehende Effizienz müssen bei der Festlegung von Effizienzvorgaben berücksichtigt werden.

Detaillierte Vorgaben zur Ausgestaltung des Verfahrens finden sich in den §§12 bis15, sowie in Anhang 3 zu § 12 der Anreizregulierungsverordnung (ARegV).

Dreh- und Angelpunkt des Effizienzvergleichs ist es, die unterschiedlichen Netzbetreiber „vergleichbar“ zu machen. Dies wird vor allem durch die Auswahl von geeigneten Vergleichsparametern gewährleistet. Vorschriften zur Auswahl von Vergleichsparametern finden sich in §13 ARegV: Hiernach sind Vergleichsparameter Parameter zur Bestimmung der Versorgungsaufgabe und der Gebietseigenschaften, insbesondere die geografischen, geologischen oder topografischen Merkmale und strukturellen Besonderheiten der Versorgungsaufgabe auf Grund des demografischen Wandels des versorgten Gebietes. Die Parameter müssen geeignet sein, die Belastbarkeit des Effizienzvergleichs zu stützen. Die Auswahl der Vergleichsparameter hat mit qualitativen, analytischen oder statistischen Methoden zu erfolgen, die dem Stand der Wissenschaft entsprechen. Durch die Auswahl der Vergleichsparameter soll die strukturelle Vergleichbarkeit der Netzbetreiber möglichst weitgehend gewährleistet werden und die Heterogenität der Aufgaben der Netzbetreiber weitgehend abgebildet werden.

Dennoch wird es Netzbetreiber geben, deren Versorgungsaufgabe nicht vergleichbar ist und auch mittels der Verwendung von Strukturparametern nicht ausreichend vergleichbar gemacht werden kann. Um zu verhindern, dass diese Unternehmen die Effizienzwerte aller Netzbetreiber beeinflussen, sind diese Unternehmen aus dem Datensatz vor Berechnung der Effizienzwerte zu entfernen. Vorschriften zur Durchführung der Ausreißeranalyse finden sich in Anlage 3 zu § 12 ARegV.

Obgleich die ARegV bereits relativ detaillierte Vorgaben zur Durchführung des Effizienzvergleichs macht, hat der Verordnungsgeber der Bundesnetzagentur dennoch einen

weitreichenden Spielraum zur Ausfüllung dieser Vorgaben gelassen. Dies wurde auch durch eine Entscheidung des Bundesgerichtshofs (EnVR 42/16) erneut bestätigt. Dieser Spielraum hat jedoch keinen Selbstzweck, sondern dient ausschließlich der Erfüllung der gesetzlichen Aufgaben aus §21a EnWG in Verbindung mit §§12 bis 15 ARegV. Gemäß dem Urteil des Bundesgerichtshofes hat die Bundesnetzagentur einen methodischen Spielraum, wie sie die Vergleichbarkeit von Netzbetreibern sicherstellt (RN 35). Wenn Hinweise bestehen, dass die strukturelle Vergleichbarkeit mit den in der Verordnung genannten Methoden (zum Beispiel zur Ausreißeridentifizierung, RN 38 BGH Beschluss) nicht gewährleistet ist, ist die Bundesnetzagentur nach dem Beschluss des Bundesgerichtshofes verpflichtet ihren Spielraum zu nutzen und weitere Anpassungen in den Berechnungsschritten vorzunehmen (RN 55).

Diese rechtlichen Voraussetzungen vorangeschickt, finden Sie in den nächsten Abschnitten unsere Anmerkungen zur Datengrundlage, Datenqualität, der Kostentreiberanalyse, Parameterpriorisierung und dem konsultierten Modell.

Da die Datengrundlage zum Zeitpunkt der Konsultation noch nicht vollständig vorliegt, behalten wir uns vor auch im Nachgang nochmals Stellung zu nehmen.

2 Datengrundlage bei Modellfindung und Konsultation

Die veröffentlichten Daten umfassen bei den Aufwandsparametern lediglich 185 der 189 VNB im Effizienzvergleich. Die vier fehlenden Netzbetreiber können eine hohe Relevanz haben und die Analyseergebnisse sowohl in der DEA als auch der SFA maßgeblich beeinflussen. So zeigt sich (Abschnitt 5.4), dass schon der Ausschluss eines einzelnen Unternehmens die Effizienzwerte aller stark beeinflussen kann.

Daher ist es aus Sicht der Netze BW dringend notwendig die Kostentreiberanalyse und Modellfindung basierend auf der gesamten Datenbasis durchzuführen. Die Modellfindung durch die Bundesnetzagentur wurde laut Konsultationsveranstaltung basierend auf 187 Datenpunkten durchgeführt und kann somit nicht von den Netzbetreibern nachvollzogen werden. Im Nachgang zur Konsultationsveranstaltung wurde zusätzlich der Modellvorschlag basierend auf den veröffentlichten 185 Datenpunkten veröffentlicht.

Die Effizienzwerte und Modelleigenschaften, die sich aus den final zu erwartenden 189 Datenpunkten ergeben, sind somit weiterhin nicht bekannt. Eine vollständige Datenbasis ist jedoch notwendig, um sich ein abschließendes Urteil über die vorgestellten Modelle bilden zu können.

Unsere Auswertungen in der vorliegenden Stellungnahme beziehen sich stets auf die uns vorliegende Datenbasis mit 185 Netzbetreibern.

Die in den Konsultationsunterlagen erwähnten Adressdichtedaten sind ebenfalls nicht im Datensatz vorhanden. Diese wurden aber, wie aus den Konsultationsunterlagen hervorgeht, im

Rahmen der Kostentreiberanalyse verwendet¹. Um diese vollständig nachvollziehen zu können ist eine Veröffentlichung der Adressdichtedaten nötig.

Die Datenveröffentlichung der in den Effizienzvergleich eingehenden Aufwands- und Strukturdaten ist in § 23b Abs. 1 S. 1 Nr. 7 EnWG vorgeschrieben.

- Die Kostentreiberanalyse muss zwingend auf dem Datensatz durchgeführt werden, welcher anschließend auch für die Berechnung der finalen Effizienzwerte verwendet wird.
- Der vollständige Datensatz muss den Netzbetreibern zur Verfügung gestellt werden.

3 Datenqualität und Eigenschaften des Datensatzes

Die Netze BW begrüßt die eingebauten Plausibilitätsprüfungen im Erhebungsbogen, die dazu geführt haben, dass sich die Datenqualität verbessert hat. Dennoch sind weitere, insbesondere ingenieurwissenschaftliche, Prüfungen sinnvoll. Bei diesen Prüfungen wäre auch ein Einbezug der ingenieurwissenschaftlichen Berater sinnvoll.

Eine Bewertung der durch die Bundesnetzagentur und den Gutachter Frontier Economics vorgenommenen Plausibilisierung ist derzeit nicht möglich, da zum jetzigen Zeitpunkt kein Gutachten veröffentlicht wurde. Lediglich das allgemeine Vorgehen, wie es im Rahmen der virtuellen Auftaktveranstaltung zum Konsultationsstart vorgestellt wurde, ist bekannt. Welche Auffälligkeiten im Zuge der Plausibilisierung identifiziert wurden und wie mit den Auffälligkeiten umgegangen wurde, ist nicht bekannt. Auch Analysen zur generellen Struktur des Datensatzes wurden in den Konsultationsfolien nicht dargestellt und können somit nicht nachvollzogen werden.

Es sind aus Sicht der Netze BW noch verschiedene Datenauffälligkeiten vorhanden, die geprüft werden müssen.

3.1 Datenauffälligkeiten:

- Es ist ersichtlich, dass es zahlreiche Netzbetreiber mit einer großen Anzahl an „Ausseipunkten im Niederdrucknetz (Betriebsdruck)“ gibt, diese aber keine Leitungslängen im Niederdrucknetz (auch keine HAL) ausweisen, dafür aber im Hochdrucknetz > 5 bar. Die Daten legen den Eindruck nahe, dass eigentliche Niederdrucknetze im Hinblick auf den Auslegungsdruck als Hochdrucknetze angegeben wurden. Es stellt sich daher die

¹ vgl. Foliensatz zur Konsultation Frontier Economics (2022), Folie 44

Frage, inwieweit dies gemäß der vorgegebenen Definition (Nr. 2 „Auslegungsdruck“: „Insbesondere beim Auslegungsdruck von Leitungen und Leitungsabschnitten ist jedoch nicht isoliert auf den Auslegungsdruck der Leitung bzw. des Leitungsabschnitts abzustellen, sondern auch auf andere im Zusammenhang stehende Anlagenkomponenten wie z.B. Armaturen, welche durch ihren Einbau den Auslegungsdruck des Leitungsabschnitts limitieren.“) korrekt ist. Im Fall einer Trennung des Parameters Leitungslänge nach Druckstufen könnten sich sonst Verzerrungen ergeben.

- Auffällig ist auch eine teilweise hohe Diskrepanz zwischen den Ausspeisepunkten >5 bar nach Betriebs- und Auslegungsdruck. Hier sollte geprüft werden, ob es sich bei den Anschlusspunkten >5 bar nach Auslegungsdruck möglicherweise um falsch zugeordnete Hausanschlüsse handelt.

3.2 Analyse des Datensatzes

3.2.1 Versorgungsaufgaben der enthaltenen Netzbetreiber

Zu den Versorgungsaufgaben gehören der Gastransport (Druckstufen > 16 bar, HD 3 und HD 4) und die Gasverteilung, die sich nochmals in Regional- (> 5 und ≤ 16 bar, HD 2) und Ortsverteilung (< 5 bar, HD 1, MD, ND) untergliedern lässt, sowie die Kombination dieser beiden Versorgungsaufgaben. Die Daten zeigen, dass einige wenige VNB in hohem Maße die Funktion des (überregionalen) Gastransport ausüben (9 VNB). Diese VNB betreiben nahezu ausschließlich Hochdruckleitungen > 16 bar, um große Durchsatzmengen über große Entfernungen zu transportieren. Darüber hinaus betreiben diese Netzbetreiber keine lokalen Ortsverteilnetze und verfügen daher über keine Konzessionsverträge und folglich über keine Konzessionsflächen.

Des Weiteren sind im Datensatz eine Vielzahl von Netzbetreibe enthalten, die eine kombinierte Versorgungsaufgabe (43 VNB) wahrnehmen. Diese Netzbetreiber zeichnen sich dadurch aus, dass sie sowohl die Versorgungsaufgabe des (überregionalen) Gastransport (> 16 bar), als auch die Regionalverteilung und die Ortsverteilung insbesondere in Mittel-/Niederdrucknetzen (auf Basis von Konzessionsverträgen) wahrnehmen.

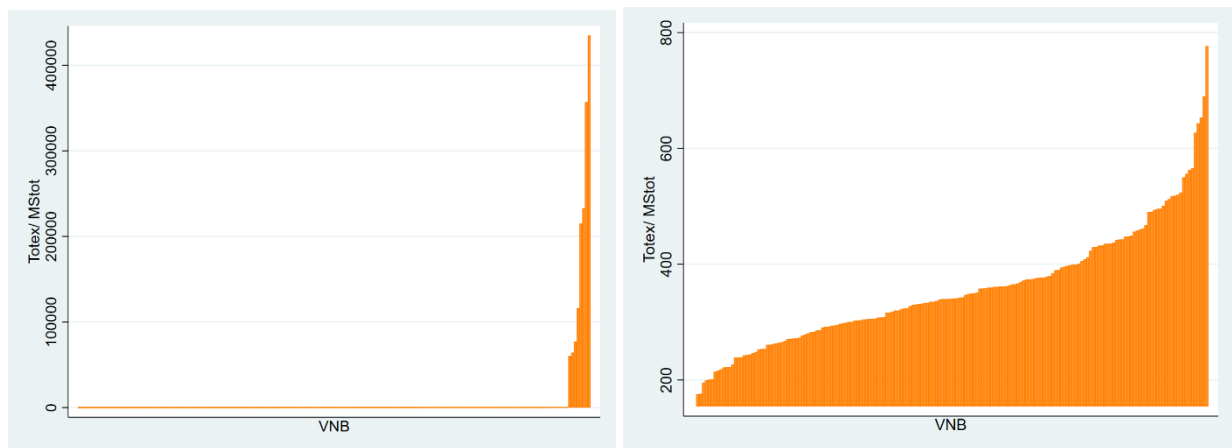
Mit Blick auf den Gesamtdatensatz ist zu konstatieren, dass der Schwerpunkt der Versorgungsaufgabe der Mehrzahl der am Effizienzvergleich teilnehmenden Netzbetreiber in der Regionalversorgung (> 5 bar und ≤ 16 bar) sowie im Ortsverteilnetz (< 5 bar) auf Basis von Konzessionsverträgen zu verorten ist (rd. 80 % aller VNB). Konkret sind im Datensatz 28 VNB enthalten, die ausschließlich in der Ortsverteilung (auf Basis von Konzessionsverträgen) aktiv sind und damit über keine Hochdrucknetze > 5 bar verfügen.

Diese unterschiedlichen Versorgungsaufgaben müssen bei der Ausgestaltung des Effizienzvergleichsmodells berücksichtigt werden.

3.2.2 Stückkostenanalyse und Kennzahlenanalyse

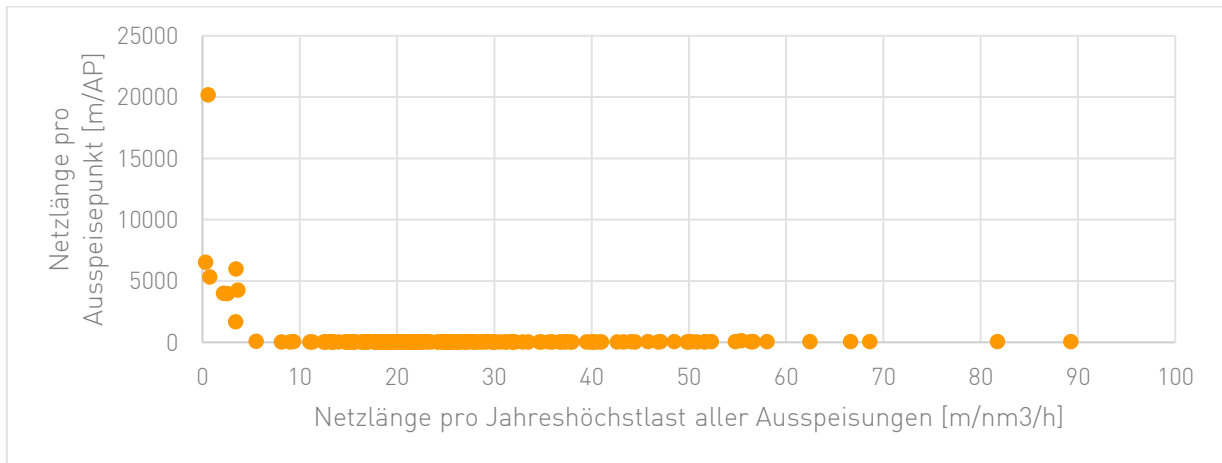
Dass der Datensatz eine sehr hohe Heterogenität aufweist, die bei der Modellfindung berücksichtigt werden muss, zeigt beispielsweise eine Analyse der Stückkosten, dem Verhältnis zwischen den Gesamtkosten eines Netzbetreibers und seiner Parameterwerte. Diese Kennzahl ermöglicht unabhängig von der Unternehmensgröße einen Vergleich zwischen den Netzbetreibern.

Exemplarisch zeigen wir in den folgenden Grafiken die Stückkosten je Messstelle. Die linke Seite der Abbildung zeigt den gesamten Datensatz. Hier fällt auf, dass es sehr wenige VNB gibt, die so extrem hohe Stückkosten aufweisen, dass für alle anderen kaum ein Unterschied sichtbar ist. Es handelt sich bei diesen wenigen VNB um die Netzbetreiber ohne Konzessionsgebiet, die nur im HD-Transportbereich unterwegs sind. Lässt man diese außen vor, ergibt sich auch für die restlichen VNB ein heterogenes Bild bei den Stückkosten (rechte Seite der Abbildung).



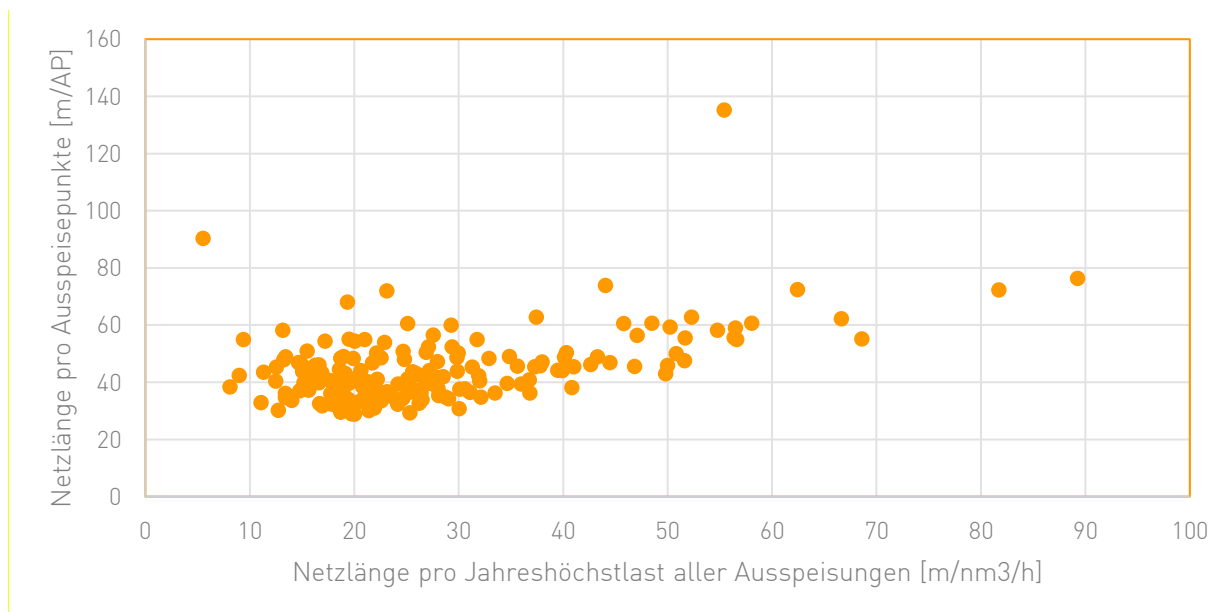
Betrachtet man eine Vielzahl von Parametern, so fällt auf, dass die Mehrzahl der Netzbetreiber an den Rändern der Verteilungen – also mit besonders geringen oder hohen Stückkosten – dadurch charakterisiert sind, dass sie entweder nur ein Hochdrucknetz betreiben oder – gemessen an den TOTEX – sehr klein bzw. sehr groß sind. Diese Netzbetreiber sollten auf das Vorliegen von Datenfehlern oder aber von strukturellen Besonderheiten überprüft werden.

Einen deutlichen Hinweis auf die Heterogenität der verglichenen Netzbetreiber zeigt auch die folgende Abbildung. Sie bildet den Zusammenhang zwischen Netzlänge pro Ausspeisepunkt und Netzlänge pro Jahreshöchstlast aller Ausspeisungen ab. Ein Unternehmen erreicht mit über 20.000 m/AP einen extrem hohen Wert. Der Großteil der Netzbetreiber weist deutlich niedrigere Werte für die Netzlänge pro Ausspeisepunkt auf und verteilt sich in der Abbildung auf der x-Achse.



Die sich vom Rest der Netzbetreiber abhebenden Netzbetreiber wurden bereits im Gutachten zum Effizienzvergleich der dritten Regulierungsperiode anhand ihrer besonderen Eigenschaften identifiziert.

Die Abbildung ändert sich deutlich, sobald die neun Netzbetreiber mit sehr hohen Netzlängen pro Ausspeisepunkt und gleichzeitig sehr niedrigen Werten für die Netzlänge pro Jahreshöchstlast aller Ausspeisungen in der Darstellung unberücksichtigt bleiben. Diese Netzbetreiber ohne Konzessionsvertrag nehmen nahezu ausschließlich eine Transportaufgabe (> 16 bar) wahr. Es wird deutlich, dass die Unternehmen innerhalb der Gruppe mit niedrigen Netzlängen pro Ausspeisepunkt ebenfalls sehr heterogen verteilt sind. Dieser Effekt wird in der vorigen Abbildung jedoch durch die Extremwerte maskiert.



Es ist offenkundig, dass hier strukturelle Unterschiede in der Versorgungsaufgabe vorliegen. Die Bundesnetzagentur sollte prüfen, in welcher Form dies bei der Ermittlung sachgerechter Effizienzwerte für alle Gasverteilernetzbetreiber zu berücksichtigen ist.

- Die Datengrundlage muss einer Reihe von Prüfungen auf ingenieurwissenschaftliche Plausibilität unterzogen werden. Die durchgeführten Prüfungen müssen dokumentiert werden.
- Die unterschiedlichen Versorgungsaufgaben der Netzbetreiber, die sich insbesondere durch die Versorgung der verschiedenen Druckstufen zeigen, müssen im Effizienzvergleich berücksichtigt werden.
- Netzbetreiber, die bei der Stückkostenanalyse auffällig sind, sollten gesondert überprüft werden.
- Mit den strukturellen Unterschieden zwischen den Netzbetreibern muss im Rahmen der Modellfindung und Ausreißeranalyse adäquat umgegangen werden.

4 Kostentreiberanalyse

Die Kostentreiberanalyse für den Effizienzvergleich Gas in der vierten Periode besteht aus drei sequenziellen Schritten:

1. Vorauswahl und Priorisierung von potenziellen Vergleichsparametern anhand ing. und statistischer Überlegungen
2. Wahl der funktionalen Form für die Stochastic Frontier Analysis (SFA)
3. Wahl der Vergleichsvariablen für das finale Modell.

Im dritten Schritt werden dabei drei Ansätze unterschieden:

- Ansatz 1: Modell der 3. Regulierungsperiode
- Ansatz 2: Ing. Prioritätenliste, wobei hier drei funktionale Formen (Translog, normiert-linear und Cobb-Douglas) separat analysiert werden.
- Ergänzend: Lasso-Methodik (ermöglicht nicht pfadabhängige Variablenselektion durch Lasso- Strafterm, der Koeffizienten bis zu einem Wert von Null schrumpfen kann)

4.1 Modellnetzanalyse und Priorisierung der Parameter entlang der Versorgungsdimensionen

Im Gegensatz zum Gutachten der dritten Regulierungsperiode, in dem die Modellnetzanalyse so gut wie gar nicht erläutert wurde, wurde in den aktuellen Konsultationsfolien auf einigen wenigen Seiten die Modellnetzanalyse dargestellt. Trotzdem reichen diese Folien ohne weitere Erläuterungen nicht dazu aus, das genaue Vorgehen der Modellnetzanalyse nachvollziehen zu können. Somit ist es uns nicht möglich die Schlussfolgerungen der Modellnetzanalyse zu überprüfen.

Auffällig ist, dass auf keiner der gezeigten Folien eine Druckstufendifferenzierung durchgeführt wird, so dass wir davon ausgehen, dass diese nicht stattgefunden hat.

Die gezeigte Analyse der kostentreibenden Wirkung der Druckstufenwahl auf Grundlage einer Betrachtung des Auslegungs- vs. Betriebsdruck von Ausspeisepunkten ist ungeeignet für eine grundlegende Bewertung des kostentreibenden Einflusses unterschiedlicher Druckstufen bzw. Netzebenen auf die Kosten von Netzbetreibern und kann als Begründung für den Verzicht auf eine Disaggregation von Parametern nicht herangezogen werden.

Insgesamt muss das Vorgehen der Modellnetzanalyse, die als Basis für Prioritätenliste der Vergleichsparameter dient, detailliert und verständlich dargelegt werden. Dies gilt auch für die Datengrundlage der Modellnetzanalyse.

Am 18.01 hat die Bundesnetzagentur den Verbänden im Nachgang eine u.a. von Herrn Prof. Dr. Müller-Kirchbauer verfasste Unterlage („Modelling the impact of the energy transition on gas distribution networks“) zur Verfügung gestellt, aus der vermeintlich die Datengrundlage der Modellnetzanalyse entnommen werden kann. Ungeachtet der Tatsache, dass die Unterlage Tabellen mit Kostendaten enthält, verwundert es, dass für den Nachweis der Datengrundlage zur Modellnetzanalyse des Effizienzvergleichs der 4.RP auf ein Papier aus dem November 2021, dass sich mit dem Einfluss der Energiewende auf die Modellierung von Gasverteilnetzen in Deutschland befasst, verwiesen wird.

Im Hinblick auf die im Papier beschriebene Datengrundlage ist anzumerken, dass die in den Tabellen auf den Seiten 10 ff. ausgewiesenen Kosten bei einem Bezug auf €/km unplausibel niedrig erscheinen (1km DN32-Leitung für 5,26€) und sich daher eigentlich nur auf €/m beziehen können. Sofern es sich dabei lediglich um ein redaktionelles Versehen handeln sollte, legt die Höhe der einzelnen Kostenbeträge je Leitungsquerschnitt aber nahe, dass es sich dabei lediglich ausschließlich um Materialkosten für die Leitungen handeln kann und die im Rahmen der Leitungsverlegung anfallenden und nicht unerheblichen Tiefbaukosten nicht berücksichtigt wurden. Sollte dies der Fall sein, stellt sich die Frage, wie im Rahmen der Modellnetzanalyse eine sachgerechte Bewertung der Kostenzusammenhänge im Hinblick auf die Dimensionierung und Ausdehnung des Netzes erfolgen kann. Dieser Sachverhalt sollte noch einmal eingehend geprüft und die final verwendeten Daten im Gutachten entsprechend dargelegt werden.

Die Modellnetzanalyse der Gutachter der Bundesnetzagentur kommt zu den gleichen Schlussfolgerungen wie in der 3. Regulierungsperiode. Es wird lediglich darauf hingewiesen, dass die

Dimension „Transport von Energie“ aufzuteilen ist in: „Verteilung von Energie an Endkunden“ und „Transport von Energie in Hochdrucknetzen“. Allerdings sind die vorgeschlagenen korrespondierenden Priorität- 1-Parameter weiterhin Messlokationen oder Ausspeisepunkte für die Verteilungsdimension und Ausspeisepunkte im Bereich >5 oder >16 bar für die Transportdimension.

Messlokationen oder Leitungslänge mit Druckdifferenzierung werden lediglich als als Prio 2 bzw. 3 eingeordnet.

Ergänzt wurde diese ingenieurwissenschaftliche Vorgehensweise durch LASSO-Regressionen mit Niveaugrößen, deren genaue Ausgestaltung jedoch den Folien nicht entnommen werden und somit nicht bewertet werden kann. Hier ist eine größere Transparenz notwendig.

- Insgesamt muss das Vorgehen der Modellnetzanalyse, die als Basis für Prioritätenliste der Vergleichsparameter dient, detailliert und verständlich dargelegt werden.

4.2 Wahl der funktionalen Form

Die Bundesnetzagentur bzw. ihre Gutachter vergleichen die vier Optionen linear, normiert-linear, logarithmiert und Translog und kommt zu dem Schluss, dass das Translog-Modell am besten geeignet ist, den Kosten-Output -Zusammenhang zu erklären.

Die Entscheidung für ein Translog-Modell wird von der Netze BW weiterhin begrüßt, da dieses Kostenzusammenhänge deutlich flexibler abbilden kann. Es sollte jedoch auch die Möglichkeit in Betracht gezogen werden, einzelne Parameter als z-Variablen ohne Interaktions- und Quadratterme zu verwenden (siehe auch 4.3 und 6).

4.3 Anzahl der Vergleichsparameter

Bei einem Translog-Modell führt schon eine geringere Anzahl an Vergleichsparametern aufgrund der Kreuz- und Quadratterme öfter dazu, dass ein Modell nicht mehr konvergiert.

Dennoch sollte diese Sorge die Anzahl der Vergleichsparameter nicht von vorneherein limitieren. Es existieren auch viele Modelle mit mehr als 5 Vergleichsparametern, die konvergieren.

Auch kann z.B. das Weglassen von Interaktionstermen für einzelne Parameter im Translog-Modell den Raum geben, die Heterogenität der Netzbetreiber durch die Hinzunahme weiterer Strukturparameter genauer abzubilden und die Modellgüte zu verbessern.

Mit den Anforderungen der Anreizregulierungsverordnung, zum einen die Heterogenität der Aufgaben von Netzbetreibern abzubilden und zum anderen möglichst nur Parameter zu verwenden, deren Wirkung ganz oder teilweise nicht wiederholend ist, entsteht zudem ein Zielkonflikt.

Hierbei ist jedoch zu berücksichtigen, dass die Abbildung korrelierter Vergleichsparameter die berechneten Effizienzwerte an sich nicht beeinflusst.²

Es sollten daher Vergleichsparameter nicht grundsätzlich aufgrund von Multikollinearität ausgeschlossen werden. Die von der Bundesnetzagentur immer wieder thematisierte Multikollinearität bei der Verwendung von mehr Vergleichsparametern stellt 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. Multikollinearität hat auch keinen Einfluss auf das Bestimmtheitsmaß R^2 (und davon abgeleitete Gütemaße) und auch nicht auf die Höhe der Residuen der Schätzung³. Hiervon abzugrenzen ist die Bedeutung der Multikollinearität bei der stufenweisen Selektion von Parametern.

4.4 Wahl der Vergleichsvariablen

Es wurden verschiedene methodische Ansätze zur Herleitung eines Modells durchgeführt. Einer dieser Ansätze war das Weiterentwickeln des Modells aus Regulierungsperiode 3, ein anderer Ansatz war die Entwicklung eines Modells durch schrittweise Variablenselektion auf Basis einer ingenieurwissenschaftlichen Prioritätenliste. Dies wurde jeweils für die funktionalen Formen Translog-, normiert- linear und Cobb- Douglas durchgeführt.

4.4.1 Ansatz 1: EVG3 Basis

Für die Weiterentwicklung des Modells aus der dritten Regulierungsperiode wird zuerst überprüft, ob dieses Modell dazu geeignet ist, den Zusammenhang zwischen Output-Parametern und Kosten für die Daten der vierten Regulierungsperiode darzustellen.⁴ Hierbei wird angemerkt, dass die Signifikanz des Ineffizienzterms mit den Daten der vierten Regulierungsperiode nicht mehr gegeben ist.

Zunächst wurde bei diesem Ansatz das Austauschen der einzelnen Modellparameter durch aus ingenieurwissenschaftlicher Sicht geeignete Parameter aus Priorität 1 und 2 geprüft.⁵ Es wird jeweils ein Parameter ersetzt während alle anderen im Modell bleiben. Verschiedene Bewertungskriterien wie die Informationsgüte, Größe/ Vorzeichen der Koeffizienten,

² Vgl. z.B. Andor, M & Hesse, F (2011), A Monte Carlo Simulation comparing DEA, SFA and two simple approaches to combine efficiency estimates, CAWM Discussion Paper No. 51, University of Münster. Jensen, U. (2005). Misspecification preferred: The sensitivity of inefficiency rankings. Journal of Productivity Analysis, 23:223-244.

³ vgl. Kennedy, P. (2008): A Guide to Econometrics 6. ed., Wiley-Blackwell, Kapitel 12.2

⁴ vgl. Foliensatz zur Konsultation Frontier Economics (2022), Folie 30

⁵ vgl. Foliensatz zur Konsultation Frontier Economics (2022), Folie 47

ingenieurwissenschaftliche Beurteilung und Konvergenz der SFA sowie Signifikanz des Ineffizienzterms werden zur Bewertung des Modells herangezogen.

Außerdem wird das Hinzufügen weiterer Parameter als zusätzliche Modellvariable getestet. Es wird allerdings in den Konsultationsunterlagen nicht spezifiziert, welche Parameter getestet wurden. Auch ob die Aufteilung von Parametern in verschiedene Druckstufen in Betracht gezogen wurde, bleibt unklar.

Dieser Ansatz ist stark pfadabhängig. Es werden zu einem vorgegebenen Modell Variablen hinzugefügt bzw. Variablen ausgetauscht. Je nachdem in welcher Reihenfolge dies geschieht, kann dies das im Ergebnis resultierende Modell beeinflussen. Bei stufenweisen Variablenselektionsmechanismen werden die Informationsgütekriterien jeweils in jedem Schritt angewendet. Diese Informationsgütekriterien sind dazu geeignet eine bestimmte Anzahl an Modellen miteinander zu vergleichen, jedoch nicht dazu diese mehrfach aufeinanderfolgend anzuwenden.⁶ Auch der t-Test verliert durch ein stufenweises Vorgehen seine Aussagekraft. Die Ursache hierfür liegt darin, dass die geschätzten Standardfehler durch die mehrfache Anwendung verzerrt sind.⁷ Es werden in jedem Schritt von allen in diesem Schritt getesteten Variablen nur die statistisch signifikanten zum Modell hinzugefügt. Dies führt auch dazu, dass bestimmte Modellkombinationen gar nicht überprüft werden. Aufgrund der verzerrten Standardfehler kann es dazu kommen, dass das Signifikanzlevel einiger Variablen als sehr niedrig eingeschätzt wird und diese Variablen im Modell enthalten sind, während andere Variablen, die einen relevanten Erklärungsgehalt besitzen, nicht in das Modell aufgenommen werden, da diese nicht als solche erkannt werden.⁸ Des Weiteren führt eine schrittweise Variablenselektion dazu, dass die Verteilungsannahme des F-Tests nicht mehr erfüllt ist und somit die Interpretation seines Ergebnisses nicht mehr möglich ist.⁹

Ein weiteres Problem, welches bei einer stufenweisen Variablenselektion vorhanden ist, ist der Einfluss von Multikollinearität. Wenn zwei oder mehr Variablen in einem Modell stark miteinander korreliert sind, wird der Effekt des einzelnen Parameters auf die abhängige Variable durch

⁶ zum Beispiel: Burnham, K.P., Anderson, D.R. & Huyvaert, K.P (2011): AIC model selection and multi-model inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* 65, 23–35

⁷ zum Beispiel: WHITTINGHAM, M.J., STEPHENS, P.A., BRADBURY, R.B. and FRECKLETON, R.P. (2006): Why do we still use stepwise modelling in ecology and behaviour?, *Journal of Animal Ecology*, 75: 1182–1189. und Smith, G. (2018): Step Away from Stepwise, *Journal of Big Data* Vol. 5, 32

⁸ zum Beispiel: Harrell, F. (2015): *Regression Modeling Strategies With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, Springer Series in Statistics S. 68. Finos, L.; Brombin, S. and Salmaso, L. (2010): Adjusting Stepwise p-Values in Generalized Linear Models, *Communications in Statistics - Theory and Methods*, 39:10, 1832–1846

⁹ zum Beispiel: Pope, P.T. and Webster, J.T. (1972): Use of an F- statistic in stepwise regression procedures, *Technometrics* Vol. 14 Nr. 2 pp. 327– 340. Harrell, F. (2015): *Regression Modeling Strategies With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, Springer Series in Statistics p. 68

den Koeffizienten nicht mehr korrekt dargestellt, auch wenn das Gesamtmodell nicht verzerrt ist, da der Gesamteffekt der korrelierten Variablen vollständig enthalten ist. Der Grund hierfür ist, dass nicht auseinandergehalten werden kann, welcher Teil der Streuung der durch beide Variablen gemeinsam erklärt wird durch welche Variable erklärt wird.¹⁰ Dadurch beeinflusst die Korrelation der beiden Variablen die Größe der Koeffizienten sowie deren Vorzeichen. Außerdem beeinflusst die Korrelation die Größe der Standardfehler und kann somit dazu führen, dass eine zum Modell hinzugefügte Variable als nicht signifikant abgelehnt wird, wenn eine starke Korrelation mit bereits im Modell enthaltenen Parametern besteht, oder dass ein Koeffizient eines bereits im Modell enthaltenen Parameters insignifikant wird.¹¹ Dies führt dazu, dass die Reihenfolge des Hinzufügens der Variablen eine große Rolle spielt. Außerdem werden für das Modell relevante Parameter unter Umständen fälschlicherweise nicht in das Modell aufgenommen, wenn die Signifikanz und die Größe/das Vorzeichen des Koeffizienten als Kriterium herangezogen werden. Laut Folie 29 der Konsultationsunterlagen werden diese Kriterien bei der Kostentreiberanalyse der Bundesnetzagentur berücksichtigt.

Generell ist es weniger problematisch einen zusätzlichen Parameter, ohne Einfluss auf die abhängige Variable einem Modell hinzuzufügen, als einen relevanten Parameter mit Einfluss auf die abhängige Variable auszulassen. Das Hinzufügen eines zusätzlichen Parameters beeinflusst die Schätzgenauigkeit des Modells nicht sondern erhöht lediglich die Varianz der Koeffizienten, während das Weglassen eines relevanten Parameters mit Einfluss auf die abhängige Variable zu einer Verzerrung der im Modell enthaltenen Koeffizienten führen kann.¹² Dies gilt auch in Bezug auf SFA und DEA.¹³

4.4.2 Ansatz 2: Ingenieurwissenschaftliche Prioritätenliste

Der von der Bundesnetzagentur als „Grüne-Wiese-Ansatz“ bezeichnete Modellfindungsprozess startet mit dem Testen von Priorität 1 Parametern zur Bildung eines Ausgangsmodells. Als nächster Schritt wurden die Bodenklassen und Grabbarkeiten einzeln hinzugefügt und getestet. Anschließend werden Priorität 2 Parameter und zuvor verworfene Parameter getestet.¹⁴

¹⁰ zum Beispiel: Derksen, S. and Keselman, H.J. (1992), Backward, forward and stepwise automate subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology*, 45: 265-282. Chatterjee, S. und A. S. Hadi (2015): *Regression analysis by example*, John Wiley & Sons

¹¹ Zum Beispiel: James, G. et. al (2013): *An Introduction to Statistical Learning with Applications in R*, Springer Science and Business Media, 2013

¹² Kennedy, P. (2008): *A Guide to Econometrics* 6. ed., Wiley-Blackwell.

¹³ zum Beispiel: Galagedera, D. U. A., und P. Silvapulle (2003): Experimental Evidence on Robustness of Data Envelopment Analysis, *Journal of the Operational Research Society* 54, 654–660.

¹⁴ vgl. Foliensatz zur Konsultation Frontier Economics (2022), Folie 68

Für die Translog- Spezifikation resultiert aus diesem Ansatz dasselbe Modell wie aus dem Ansatz 1, der auf dem Modell der dritten Regulierungsperiode aufsetzt.

Auch hier bleibt unklar, welche Variablenkombinationen in welcher Reihenfolge getestet wurden. Außerdem wurden wie auf Folie 68¹⁵ des von den Gutachtern vorgelegten Foliensatzes beschrieben die Parameter der Priorität 3 hierbei ausgelassen, obwohl die Adressdaten, welche der Priorität 3 angehören, im Ansatz zur Weiterentwicklung des Modells der dritten Regulierungsperiode getestet wurden und teilweise eine hohe Korrelation der Priorität 3 Parameter mit den Aufwandsparemtern besteht¹⁶. Des Weiteren ist unklar, ob auch disaggregierte Parameter berücksichtigt wurden und ob das Ersetzen von im Modell enthaltenen Parametern durch alternative Parameter, welche dieselbe Dimension der Versorgungsaufgabe oder auch verschiedene Dimensionen der Versorgungsaufgabe repräsentieren, getestet wurde (beispielsweise das Ersetzen des Parameters der vorherrschenden Bodenklasse NIV2BK456 durch die Grabbarkeit).

Um die einzelnen Entscheidungen für jeden Schritt bezüglich der Aufnahme von Variablen in das Modell nachvollziehen zu können, ist die Darstellung der Ergebnisse - sowohl statistische Größen als auch die ingenieurwissenschaftliche Beurteilung - für jede der getesteten Variablen nötig.

Wie im Abschnitt zum Ansatz der Weiterentwicklung des Modells der dritten Regulierungsperiode aufgeführt, hat solch ein stufenweiser Variablenselektionsansatz einige Nachteile. Vor allem die Sensitivität gegenüber der Reihenfolge, in welcher die Variablen zum Modell hinzugefügt werden, spielt auch hier wieder eine bedeutende Rolle.

4.4.3 Ansatz 3: Lasso als Ergänzung zur Modellfindung:

In den Unterlagen der Gutachter der Bundesnetzagentur wird die Lasso-Methode als ergänzender Ansatz für den Modellfindungsprozess genannt. Jedoch wird nicht näher auf diesen Ansatz und das genaue Vorgehen eingegangen. Es ist lediglich das Ergebnis, dass die resultierenden Modelle auf Grund von Überspezifikation nicht für die Anwendung im Effizienzvergleich geeignet sind, aufgeführt.¹⁷ Eine detaillierte Beschreibung der Vorgehensweise fehlt in den Unterlagen zur Konsultation, sodass dieses Ergebnis nicht überprüft bzw. nachvollzogen werden kann. Es wäre außerdem sinnvoll die in den resultierenden Lasso-Modellen enthaltenen Variablen mit denen der Modelle aus den anderen Variablenselektionsmechanismen zu vergleichen.

¹⁵ vgl. Foliensatz zur Konsultation Frontier Economics (2022), Folie 68

¹⁶ vgl. Foliensatz zur Konsultation Frontier Economics (2022), Folie 33

¹⁷ vgl. Foliensatz zur Konsultation Frontier Economics (2022), Folie 30

- Die Vorgehensweise bei der Kostentreiberanalyse muss transparent und nachvollziehbar dargestellt werden, sodass ein Nachrechnen der Ergebnisse möglich ist.
- Ein Vergleich der Ergebnisse des Lasso- Verfahrens mit den Ergebnissen der anderen Modellfindungsverfahren auf Überschneidungen und Unterschiede in Bezug auf die ausgewählten Parameter sollte durchgeführt werden.
- Um die Pfadabhängigkeit zu berücksichtigen, sollten zusätzlich zu pfadabhängigen Variablenselektionsmechanismen andere Vorgehensweisen, wie eine Änderung der Reihenfolge des Hinzufügens der Parameter oder alternative Modellfindungsverfahren durchgeführt werden. Die Ergebnisse sind zu vergleichen, um sicherzustellen, dass alle relevanten Parameter im Modell enthalten sind.

4.5 Nach Druckstufen differenzierte Betrachtung fehlt

Auf eine nach Druckstufen differenzierte Betrachtung wird im Foliensatz nur an einigen wenigen Stellen und im Wesentlichen nur bei den Ausspeisepunkten eingegangen.

Es sollte jedoch ausführlich überprüft werden, ob bestimmte Sachverhalte in aggregierter oder disaggregierter Form abzubilden sind. Parameteraggregationen können zu stark verzerrten Effizienzwerten führen, wenn die Kostenwirkungen der Parameter unterschiedlich sind.¹⁸

Während im Strom-Effizienzvergleich schon längst etabliert ist, dass verschiedene funktionale Netzebenen bei der Leitungslänge oder der installierten Leistungen mit verschiedenen hohen Kosten einhergehen, wurde eine Differenzierung im Gas-Bereich nach Druckstufen von der Bundesnetzagentur bislang also endogen bezeichnet und abgelehnt. In den aktuellen Konsultationsunterlagen finden sich lediglich unter Priorität 2 und 3 auch disaggregierte Parameter nach Druckstufen.

Die Gliederung der Netze in funktionale Netzebenen (Ferntransport (FNB), überregionaler Transport, regionale Verteilung und Ortsverteilung) ist jedoch durchaus exogen vorgegeben. Es wäre technisch und ökonomisch abwegig, etwa eine Wohnsiedlung mit den für den überregionalen Transport üblichen Drücken und Leitungstypen zu erschließen, und es wäre praktisch unmöglich, die regionale Verteilung oder den überregionalen Transport mit Niederdruckleitungen zu bewerkstelligen.

¹⁸ Vgl. Fare, R., Grosskopf, S. & Lovell, C. A. K. (1988), An indirect approach to the evaluation of producer performance, *Journal of Public Economics*, 37: 71-89.

Zwei Netze mit gleichem Rohrvolumen können beispielsweise unterschiedliche Kosten verursachen, allein weil der Netzbestand auf unterschiedliche (exogen) Netzebenen (und somit Druckstufen) und Leitungslängen verteilt sein kann. Das Rohrvolumen allein kann solche Unterschiede nicht abbilden.

Hier bietet sich beispielsweise eine Aufteilung der Netzlängen nach Druckstufenklassen an, um diese Unterschiede im Effizienzvergleich abzubilden.

Die Verwendung von disaggregierten Variablen erhöht die Anzahl der verwendeten Parameter und kann die Konvergenz der Modelle in der SFA erschweren. Dennoch gibt es zahlreiche Modelle, die auch als Translog Modell mit mehr als fünf Parametern konvergieren. Zudem ist eine Verwendung eines vollständigen Translog-Modells nicht notwendig. Durch Verwendung einzelner Parameter als z-Variablen (ohne Interaktionsterme) kann die Gesamtzahl der Modellparameter gesenkt werden. Damit wird Konvergenz erleichtert und es können bessere Gütemaße erzielt werden, da auf unnötige Terme verzichtet wird. Diese Möglichkeiten erweitern den Raum an potenziellen Effizienzvergleichsmodellen um ein Vielfaches und sollten nicht außer Acht gelassen werden. So kann der Heterogenität der Netzbetreiber besser Rechnung getragen werden, ohne die Berechnung unnötig zu überladen.

- Bei der Kostentreiberanalyse muss die disaggregierte Betrachtung nach Druckstufen genauer geprüft werden, da die Druckstufenwahl aufgrund der Versorgungsaufgabe stark exogen geprägt ist.
- Die Problematik der Nicht-Konvergenz in der SFA durch zu viele Vergleichsparameter kann dahingehend vermieden werden, dass in der Translog-Funktion einzelne Parameter ohne Interaktionsterm verwendet werden.

5 Beurteilung des konsultierten Modells

5.1 Anwendbarkeit des konsultierten Modells

Die Bundesnetzagentur verwendet im Ausgangspunkt der Analysen das Modell der 3. Regulierungsperiode mit den folgenden fünf Strukturparametern: Jahreshöchstlast der Ausspeisungen, Anschlusspunkte >5 bar, Messtellen, Rohrvolumen und Netzlänge in den Bodenklassen 4,5,6. In 0-1m Tiefe.

Das Modell erfüllt einen Großteil der statistischen Kriterien, die zur Modellbewertung durch die Bundesnetzagentur herangezogen werden. So konvergiert das Modell, ist nicht von Heteroskedastizität belastet, weist positive und signifikante Koeffizienten bei den Hauptvariablen auf und

besteht die Tests zur Modellspezifikation - den Link-Test und den Reset Test. Auch weisen die Residuen der OLS-Schätzung die nötige Rechtsschiefe auf.

Aufgrund der Ergebnisse des LR-Tests auf Signifikanz des Ineffizienzterms, die darauf hinweisen, dass aus statistischer Sicht der Fehlerterm nicht mit hinreichender Genauigkeit in Ineffizienz und „Rauschen“ unterteilt werden kann, nimmt die Bundesnetzagentur jedoch eine Anpassung vor.

Sie ersetzt den Parameter „Netzlänge in den Bodenklassen 4,5,6 in 0-1 m Tiefe“ durch den Parameter „Netzlänge in den Bodenklassen 4,5,6 in 0-2 m Tiefe“. Diese Anpassung führt dazu, dass die statistischen Kriterien inklusive des LR-Tests auf Signifikanz des Ineffizienzterms erfüllt sind.

5.2 Anpassung im Vergleich zur dritten Regulierungsperiode

Der Austausch des Parameters „Netzlänge in den Bodenklassen 4,5,6 in 0-1 m Tiefe“ durch den Parameter „Netzlänge in den Bodenklassen 4,5,6 in 0-2 m Tiefe“ führt dazu, dass das Modell die gängigen Kriterien der Bundesnetzagentur erfüllt. Dieser neue Parameter wird gebildet, indem die Flächen in den Tiefenstufen 0-1 und 1-2 addiert werden und dann durch die 2-fache Gesamtfläche geteilt werden. Dieser Wert wird anschließend mit der Netzlänge multipliziert.

Es ist jedoch aus dem bisherigen Vortag der Bundesnetzagentur nicht nachvollziehbar, auf welchen ingenieurwissenschaftlichen Überlegungen dieser Parametertausch beruht. Vielmehr scheint hier ein Vorgehen gewählt worden zu sein, bei dem nach einer Adaption des Modells der dritten Regulierungsperiode gesucht wurde, das „passende“ p-Werte beim LR-Test auf Signifikanz des Ineffizienzterms liefert.

Die Tatsache, dass bei unveränderter Anwendung des Modells der dritten Regulierungsperiode beim LR-Test auf Signifikanz des Ineffizienzterms das gängige Signifikanzniveau überschritten wird, sollte aber nicht grundsätzlich dazu führen, einen aus ingenieurwissenschaftlicher Sicht sinnvollen Parameter zu verwerfen. Dies wird auch an anderer Stelle von der Bundesnetzagentur so gehandhabt. So konnten wir beispielsweise feststellen, dass das Kriterium der Signifikanz des Ineffizienzterms bei den Malmquist-Berechnungen für den Xgen gar keine Rolle gespielt hat und in den Berechnungen regelmäßig insignifikante Ineffizienzterme resultierten.

Vielmehr werfen die hohen p-Werte zunächst einmal die Frage auf, ob Ineffizienz überhaupt noch gegeben ist. In einem solchen Fall der Unsicherheit sollte vermieden werden zu hohen Effizienzanforderungen zu stellen., Netze BW würde daher vorschlagen zumindest die SFA-Effizienzwerte so zu skalieren, dass der effizienteste Netzbetreiber (Ausreißer ausgenommen) einen Effizienzwert von 100% erreicht.

- Der Austausch des Parameters Netzlänge in den Bodenklassen 4,5,6 in 0-1 m Tiefe“ durch den Parameter „Netzlänge in den Bodenklassen 4,5,6 in 0-2 m Tiefe“ muss ingenieurwissenschaftlich begründet werden.
- Statt nach „passenden“ Parametern zu suchen, sollte die Nicht-Signifikanz des Ineffizienzterms als Hinweis gewertet werden, dass Ineffizienzen abgebaut wurden.
- In einem solchen Fall der Unsicherheit sollte daher der Fokus daraufgelegt werden, keine zu hohen Effizienzanforderungen zu stellen

5.3 Monoton steigende Kostenfunktion

Es ist zu prüfen, ob die theoretische Annahme von monoton steigenden Kosten, das heißt, ob der Anstieg eines Outputs mit steigenden effizienten Kosten einhergeht, erfüllt ist. Wie in der dritten Regulierungsperiode gibt es auch hier, beispielsweise bei den Anschlusspunkten >5 bar, Fälle, bei denen eine Erhöhung des Parameters für einen Netzbetreiber zu einer Verringerung seines Effizienzwerts führt. Dies sollte einer genaueren Analyse unterzogen werden. Argumentationen der Bundesnetzagentur, dass hier Tests zur statistischen Signifikanz durchgeführt wurden (siehe Gutachten zur 3. Regulierungsperiode), reichen nicht aus, eine solche nicht-monoton steigende Kostenfunktion zu rechtfertigen. Relevant für die betroffenen Netzbetreiber ist, dass sich ihr Effizienzwert „in die falsche Richtung“ ändert, wenn ein Parameter ansteigt.

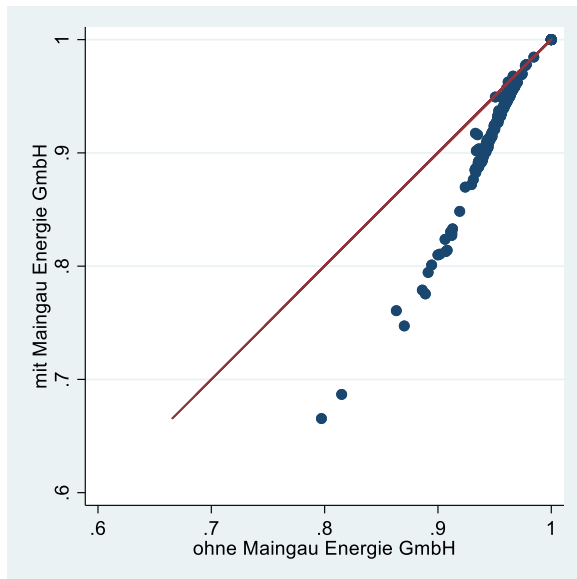
5.4 Ausreißeranalyse in der SFA

Bislang wurde in der SFA lediglich eine Ausreißeranalyse auf Basis von Cook's Distance-Werten durchgeführt. Diese Vorgehensweise testet den Einfluss eines Unternehmens auf die durchschnittliche Kostengrenze („fitted values“), berücksichtigt allerdings nicht, dass einzelne Netzbetreiber auch einen wichtigen Einfluss auf die weiteren zu schätzenden Verteilparameter der SFA haben, welche ebenfalls zur Bestimmung der effizienten Kostengrenze relevant sind. Es sollte daher untersucht werden, ob auffällige Netzbetreiber diese Parameter (wie z.B. die Standardabweichung des Ineffizienzterms) in der SFA beeinflussen. Dies ist z.B. mittels DFBETAS.

In der Literatur ist bekannt, dass Unternehmen mit auffallend geringem Effizienzwert die Effizienzwerte aller anderen Unternehmen stark beeinflussen können.

Vorliegend ist das auch hier der Fall. Um diese Problematik aufzuzeigen, haben wir das ineffizienteste Unternehmen ohne Ausreißerstatus (Maingau Energie GmbH) aus dem Datensatz entfernt und die resultierenden bestabgerechneten Effizienzwerte den Effizienzwerten gegenübergestellt, bei denen dieses Unternehmen nicht entfernt wird. Der Ausschluss dieses einzelnen VNB verbessert die Effizienzwerte aller anderen VNB maßgeblich und führt zu einer deutlich höheren Durchschnittseffizienz (Erhöhung um mehr als 3 Prozent)

Dennoch wurde dieses Unternehmen nicht als Ausreißer identifiziert. Die Tatsache, dass dieses Unternehmen von der Ausreißeranalyse auf Basis von Cook's Distance Werten nicht identifiziert wurde, kann nur bedeuten, dass dieses Unternehmen den Verteilparameter des Ineffizienzterms beeinflusst. Dies spricht sehr nachdrücklich dafür, dass auch der Einfluss einzelner Netzbetreiber auf den Ineffizienzterm in der Ausreißeranalyse berücksichtigt werden sollte.



Das identifizierte Unternehmen kann über eine Anpassung der Ausreißeranalyse als weiterer Ausreißer identifiziert werden. Ermittelt man die DFBETA-Werte für den Verteilparameter des Ineffizienzterms σ_u , so zeigt sich ein DFBETA Wert bei TOTEX von 1.39 - deutlich höher als bei allen anderen VNB. Bei den STOTEX ergibt sich ein DFBETA-Wert von 0.11., wobei auffällig ist, dass die Werte aller anderen VNB negativ sind, was für eine besondere Stellung des VNB spricht.

Wendet man nun den gängigen Grenzwert bei der Ermittlung von DFBETAs von 1 an, so wird die Maingau Energie GmbH bei TOTEX als weiterer VNB als Ausreißer erkannt. Dies verbessert den Effizienzwert aller Netzbetreiber.

Es liegt somit ein in der ARegV namentlich genanntes Verfahren vor, das den Einfluss auf die Parameter des Modells (hier σ_u) messen kann und einen weiteren Ausreißer identifiziert, der auch nachweislich großen Einfluss auf die Effizienzwerte der anderen VNB hat.

Die Netze BW regt an dieses Verfahren anzuwenden, um den starken Einfluss einzelner VNB zu begrenzen.

- Neben der Anwendung der Cook's Distance sollte ein Verfahren angewandt werden, dass den Einfluss einzelner VNB auf die Lageparameter der Ineffizienzkomponente misst.
- So kann verhindert werden, dass einzelne Netzbetreiber einen substantiellen Einfluss auf die Effizienzwerte aller anderen VNB haben.
- Hier bietet sich die Ermittlung von DFBETA-Werten für den Verteilparameter σ_u an.

5.5 SFA -Werte von 100% ermöglichen

Der maximal erreichbare SFA-Effizienzwert liegt im konsultierten Modell bei ca. 97,7%. Dies liegt darin begründet, dass ein logarithmiertes Modell im Vergleich zu einem normiert-linearen Modell auch bei vergleichbarer numerisch ermittelter Ineffizienz zu deutlich geringeren Effizienzwerten führt. Dies ist durch die Tatsache bestimmt, dass bei normierten Modellen auch die absolute Höhe der effizienten Kosten eine Rolle für den Effizienzwert spielt, während beim Translog-Modell lediglich die numerische Ineffizienz eingeht. Dieser methodisch bedingte Effekt sollte behoben werden, so dass bei einer Überführung der numerischen Ineffizienzen in eine Effizienzvorgabe sichergestellt wird, dass Effizienzwerte von 100% auch in der SFA erreichbar sein können.

Auch aufgrund der Tatsache, dass nur ca. 17% der Netzbetreiber ihren bestabgerechneten Effizienzwert aus der DEA erhalten, und diese somit zunehmend an Bedeutung verliert, ist eine Erreichbarkeit von 100% in der SFA dringend geboten.

Dies ist möglich, indem man den Effizienzwert des effizientesten Unternehmens (exklusive Ausreißer) auf 100% setzt und die anderen Werte dementsprechend hochskaliert. Die Ausreißer erhalten in diesem Fall dann maximal 100%, analog zur DEA.

- Die Effizienzwerte in der SFA müssen auf 100% hochskaliert werden, um den methodenimmanenten „gläsernen Deckel“ zu beheben

5.6 Parametrisierung in der DEA

Dass die DEA-Methode bei der Abbildung von Heterogenität größere Schwierigkeiten aufweist als die SFA, zeigte sich bereits daran, dass nur 18% der Netzbetreiber in der dritten Regulierungsperiode ihren Effizienzwert aus der DEA erhalten haben. Mit den Daten der 4. Regulierungsperiode und dem konsultierten Modell wiederholt sich diese Problematik, so dass nur noch ca. 17% der Netzbetreiber ihren Effizienzwert aus der DEA erhalten. Die Outputgewichte in

der DEA zeigen, dass die Vergleichsparameter "Jahreshöchstlast" und "Rohrvolumen" für die meisten VNB nur ein geringes Gewicht aufweisen, Die meisten VNB werden von den ehemaligen regionalen FNBs mit ihren extrem günstigen Kosten in Bezug auf diese Versorgungsdimensionen so stark dominiert, dass diese Parameter kaum in den Effizienzwert einfließen. Somit bleiben für die meisten VNB nur maximal die 3 restlichen Parameter Messtellen, Anschlusspunkte >5 bar und Netzlänge in den Bodenklassen 4,5,6 als Treiber des Effizienzwerts. Dies ist nicht ausreichend, um die komplette Versorgungsaufgabe eines VNB abzubilden.

In der Konsequenz ist zu prüfen wie die DEA-Berechnung um weitere Vergleichsparameter, die der Heterogenität der Netzbetreiber gerecht werden, zu ergänzen ist.

Die Bundesnetzagentur und ihre Gutachter sind bislang der Auffassung, dass nach den Vorgaben der ARegV in der DEA und SFA zwingend die gleichen Parameter abgebildet werden müssen. Dies schränkt die Anwendung beider Methoden jedoch unnötig ein und wird den Eigenschaften beider Methoden nicht gerecht.

Die SFA lässt (Modell-)Fehler bis zu einem gewissen Umfang zu und kann fehlende Kostentreiber zumindest noch teilweise verkraften. Die DEA als deterministisches Verfahren kann dies gerade nicht. Hier ist es notwendig, dass ein Modell komplett ist. D. h. vor dem Hintergrund der gemäß ARegV besonders zu berücksichtigenden Heterogenität ist eine ausreichende Anzahl an Parametern zu verwenden. Die Anzahl der in der DEA verwendeten Parameter sollte nicht nur aufgrund einer Limitierung des SFA-Modells beschränkt werden. Insbesondere für die DEA wird in der wissenschaftlichen Literatur das Verzerrungspotential eines fehlerhaft nicht aufgenommenen, aber relevanten Vergleichsparameters wesentlich höher eingeschätzt als das Verzerrungspotential eines fehlerhaft aufgenommenen irrelevanten Vergleichsparameters¹⁹. Vereinfacht ausgedrückt: Durch einen fehlerhaft aufgenommenen, aber irrelevanten Kostentreiber werden – wenn überhaupt – die Effizienzwerte von wenigen Unternehmen überschätzt. Durch einen fehlerhaft nicht aufgenommenen, aber relevanten Kostentreiber werden die Effizienzwerte vieler Unternehmen unterschätzt. Dieses Risiko von statistischen Fehlern erster und zweiter Ordnung sollte gegeneinander aufgewogen werden. Im Zweifelsfalle (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. Mit der Translog-Spezifikation konvergieren in einer von uns angelegten breiten Kostentreiberanalyse gerade einmal ca. 50% (bei TOTEX) der Szenarien. Nach dem bisherigen Vorgehen der Bundesnetzagentur würden nur diese Szenarien für die DEA in Erwägung gezogen werden.

Nach der Auffassung der Netze BW ist eine solch starke Einschränkung aus rechtlicher Sicht nicht notwendig: Es ist nach § 12 Abs. 4a ARegV zwar vorgeschrieben, dass jeweils die gleichen

¹⁹ Zum Beispiel Smith, P. (1997), Model Misspecification in Data Envelopment Analysis, *Annals of Operations Research*, 73(1): 233-252. Galagedera, D. & Silvapulle, P. (2003), Experimental evidence on robustness of data envelope analysis, *Journal of the Operational Research Society*, 54: S. 654-660

Vergleichsparameter verwendet werden müssen, wenn die Berechnungen für die verschiedenen Aufwandsparameter (Kosten nach StromNEV oder unter Berücksichtigung der Vergleichbarkeitsrechnung) durchgeführt werden. Eine ähnliche Einschränkung sieht die Verordnung für die Berechnung von Effizienzwerten nach den verschiedenen Methoden (SFA und DEA) nicht vor. Ein von uns in Auftrag gegebenes Rechtsgutachten zu dieser Thematik, das der Bundesnetzagentur ebenfalls vorliegt, kommt zu folgendem Ergebnis:

„DEA und SFA sind zwei gleichwertige Methoden. Die Auswahl der Vergleichsparameter für die DEA und die SFA hat mit qualitativen, analytischen oder statistischen Methoden zu erfolgen, die dem Stand der Wissenschaft entsprechen, § 13 Abs. 3 S. 7 ARegV. Diese Vorgabe ist durch die Regulierungsbehörde unabhängig davon zu beachten, dass ihr im Hinblick auf die Auswahl der Vergleichsparameter ein Ermessensspielraum zusteht.

Nach der Rechtsprechung des BVerfG sowie das BVerwG ist der Begriff „Stand der Wissenschaft“ umfassend zu verstehen. Wenn das Gesetz daher die Berücksichtigung des Standes der Wissenschaft fordert, sind alle vertretbaren wissenschaftlichen Erkenntnisse in Erwägung zu ziehen.

Die Regulierungsbehörde hält sich daher nur dann an den ihr eröffneten Spielraum, wenn sie in ihrer Bewertung all diejenigen Vergleichsparameterkombinationen berücksichtigt, die im Rahmen einer vertretbaren wissenschaftlichen Herangehensweise in Betracht kommen. Diesen Anforderungen genügt es nicht, wenn ohne unterschiedliche Parametrisierungen zu prüfen, die identischen Parameter für SFA und DEA herangezogen werden und dadurch Vergleichsparameter, die unter Umständen genauere Ergebnisse liefern würden, von vornherein nicht in die Betrachtung einbezogen werden.“

(Rechtsgutachten von White und Case vom 22.01.2019, S. 13/14)

Während die Bundesnetzagentur seit der dritten Regulierungsperiode bei der Wahl der funktionalen Form in der SFA nun auch flexiblere Formen als normierte Modelle verwendet, u.a. mit Interaktionstermen zwischen Vergleichsparametern, ist sie aufgrund ihrer rechtlichen Bedenken noch nicht bereit gewesen gänzlich unterschiedliche Parameter in DEA und SFA zu prüfen. Dieser Schritt wäre aus unserer Sicht jedoch notwendig, um der Anforderung aus der Anreizregulierung, nämlich die Heterogenität der Aufgaben der Netzbetreiber möglichst weitgehend abzubilden, gerecht zu werden.

Die vor kurzem ergangene Entscheidung des OLG Düsseldorf zum Effizienzvergleich 5 Kart 3/21 legt klar dar, dass eine separate Kostentreiberermittlung und/oder die Verwendung unterschiedlicher Vergleichsparameter durch die Vorgaben der ARegV nicht ausgeschlossen sind. Dies bedeutet auch: Wenn die Heranziehung identischer Kostentreiber ungeeignet oder ein anderes methodisches Vorgehen ihr deutlich überlegen ist, kann eine separate Parametrierung für DEA und SFA verwendet werden.

- Eine getrennte Parametrisierung bei DEA und SFA - mit dem Ziel die Heterogenität der VNB in der DEA besser abzubilden - ist zu prüfen.

5.7 Anpassungsbedarf: Ausreißeranalyse in der DEA

In der Vergangenheit wurde in der DEA eine Dominanzanalyse auf Basis eines F-Tests durchgeführt. Dieser F-Test ist aus zwei Gründen nicht anwendbar: Zum einen stammen die Effizienzwerte aus einer nicht-parametrischen Analyse. Parametrische Testverfahren sollten daher nicht angewendet werden. Zum anderen geht der Test von der Annahme aus, dass die beiden zu vergleichenden Werte aus zwei unabhängigen Stichproben stammen.²⁰ Diese Annahme ist im vorliegenden Fall klar verletzt, denn es werden zwei verschiedene Effizienzwerte des gleichen Unternehmens miteinander verglichen. Die Dominanzanalyse sollte daher auf Basis nicht-parametrischer Tests durchgeführt werden, welche die „paired“ Struktur der vorliegenden Daten berücksichtigt (z.B. Wilcoxon-Test, bootstrap test).

Es ist zu befürchten, dass wie in der dritten Regulierungsperiode durch die Hinzunahme der ehemaligen regionalen Fernnetzbetreiber Ausreißer verdeckt werden (in der Literatur als „Masking“ bezeichnet). Die Branche hat in der Vergangenheit ein wiederholtes Vorgehen der Supereffizienz-Ausreißeranalyse angeregt. Die Bundesnetzagentur befürchtet, dass dieses Vorgehen – ähnlich dem Schälen einer Zwiebel – zu viele Ausreißer identifiziert. Das Risiko von „Masking“ sollte jedoch nicht ohne eine weitere Untersuchung des Sachverhalts generell ausgeschlossen werden. So wurden in der dritten Regulierungsperiode – anders als in der zweiten Regulierungsperiode – nicht alle ehemaligen rFNB durch die Ausreißeranalyse erkannt. Dies zeigt sich auch mit den Daten der vierten Regulierungsperiode.

Es ergeben sich im konsultierten Modell selbst nach Ausschluss der Ausreißer noch Supereffizienzwerte von 127%. Das heißt diese VNB könnten ihre Kosten um 27% erhöhen und wären immer noch 100% effizient. Dies ist realitätsfern und zeigt vielmehr, dass die DEA-Berechnung mit den gegebenen Vergleichsparametern um weitere Ausreißer bereinigt werden muss. Eine wiederholte Supereffizienzanalyse ist hier angezeigt.

²⁰ Vgl. Banker, R. (1996), Hypothesis Tests Using Data Envelopment Analysis, *Journal of Productivity Analysis*, Vol. 7; 142

- Die Dominanzanalyse sollte auf Basis nicht-parametrischer Tests durchgeführt werden, welche die „paired“ Struktur der vorliegenden Daten berücksichtigt (z.B. Wilcoxon-Test, bootstrap test).
- Eine wiederholte Supereffizienzanalyse ist angezeigt.

6 Beispiel für ein alternatives Modell mit Druckstufendifferenzierung

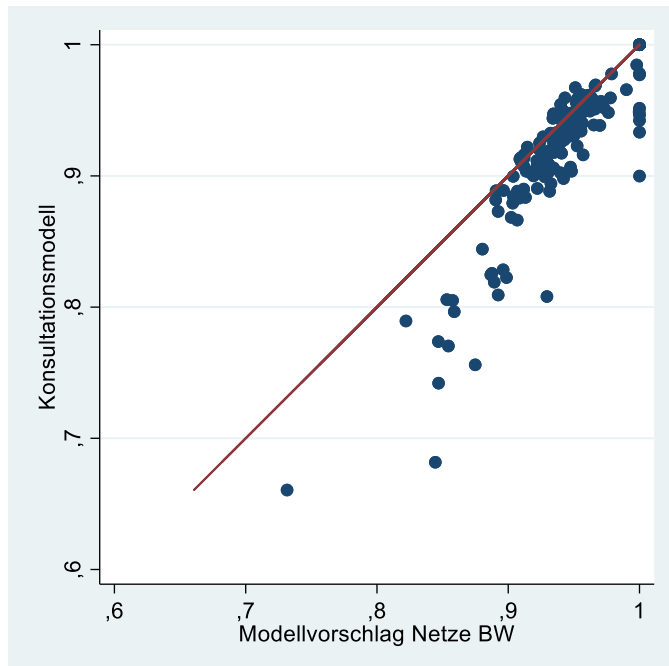
Wie in den vorangegangenen Abschnitten bereits erläutert, sollte eine druckstufendifferenzierte Betrachtung z.B. der Leitungslänge geprüft werden, um die Heterogenität der Versorgungsaufgabe der Netzbetreiber besser abzubilden. Im Folgenden zeigen wir eine von sicher einigen Möglichkeiten für ein druckstufendifferenziertes Modell.

Wir verwenden als Vergleichsparameter die Jahreshöchstlast, Messstellen, Rohrvolumen, Anschlusspunkte > 5 bar wie im konsultierten Modell; zusätzlich wird der Netzlänge-Parameter in den Bodenklassen 4,5,6 in Tiefenstufe 0-2 Meter (analog Konsultation) aufgeteilt auf drei Parameter, die jeweils die Druckstufen <5 bar, 5-16 bar und größer 16 bar abbilden und somit die verschiedenen grundsätzlichen Versorgungsdimensionen der VNB (Ortsverteilung, Regionaltransport, Ferntransport) darstellen. Diese Parameter werden in der SFA jedoch als z-Variablen verwendet und nicht mit dem Rest interagiert, um das Modell nicht stark zu vergrößern.

In der SFA konvergiert das Modell, die Hauptkoeffizienten haben positive Vorzeichen und sind (bis auf Netzlänge in BK 4,5,6 >16 bar) signifikant. Heteroskedastizität kann abgelehnt werden und das korrigierte BIC bewegt sich in gleicher Höhe wie im konsultierten Modell. Die Ineffizienzterme sind nicht signifikant, was analog zur Anwendung des Modells der dritten Regulierungsperiode mit den Daten der vierten Regulierungsperiode ein Hinweis auf nicht (mehr) in hohem Maße vorhandene Ineffizienzen ist, aber – so die Sicht der Netze BW- nicht dazu führen sollte ein aus ingenieurwissenschaftlicher Sicht sinnvolles Modell grundsätzlich zu verwerfen.

In der DEA erhalten deutlich mehr VNB ihren Effizienzwert aus der DEA (ca. 25%), so dass die DEA insgesamt mehr Gewicht erhält.

Die bestabgerechnete Effizienzwerte im Vergleich zum Konsultationsmodell werden in folgender Grafik dargestellt. Der durchschnittliche Effizienzwert nimmt zu, die Standardabweichung sinkt. Insgesamt profitieren ein Großteil der Netzbetreiber deutlich von dieser Anpassung.



7 Plausibilisierung der Ergebnisse

Aus den Konsultationsunterlagen ist nicht ersichtlich ob und wie die Ergebnisse plausibilisiert werden. In der Vergangenheit wurde zur Modellvalidierung eine Second Stage Analyse angewendet. Bei einer Second Stage wird mittels graphischer Darstellungen und (Tobit) Regressionen überprüft, ob ein systematischer Zusammenhang zwischen den bestabgerechneten Effizienzwerten eines Modells und ausgelassenen Vergleichsparametern besteht. Dieses Vorgehen kann aus Sicht der Netze BW nicht angewendet werden: Grundsätzlich wurde die Second Stage Analyse nicht zur Modellvalidierung entwickelt und ist in der DEA – wenn überhaupt – nur für Umweltparameter zulässig²¹ und in der SFA grundsätzlich nicht anwendbar.²² Der Einfluss von im Modell unberücksichtigten Parametern auf die Effizienzwerte kann mit Hilfe der Second Stage Analyse nicht beurteilt werden.

²¹ vgl. Banker, R.D. & Natarajan, R. (2008), Evaluating contextual variables affecting productivity using data envelopment analysis, *Operations Research*, Vol. 56(1), 48-58; Simar, L. & Wilson, P.W. (2007), Estimation and inference in two-stage, semi-parametric models of production processes, *Journal of Econometrics*, 136(1): S. 31-64; Simar, L. & Wilson, P.W. (2011), Two-stage DEA: caveat emptor, *Journal of Productivity Analysis*; Vol. 36: S. 205; Johnson, A.L. & Kuosmanen, T. (2012), One-stage and two-stage DEA estimation of the effects of contextual variables, *European Journal of Operational Research*, Vol. 220(2): S. 559-570.

²² vgl. Wang, H.J. & Schmidt, P. (2002), One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels, *Journal of Productivity Analysis*, 18(2): S. 129-144.

Aus diesen Gründen sollte auf eine Anwendung der Second-Stage-Analyse grundsätzlich verzichtet werden und geeignete Verfahren zur Modellvalidierung angewendet werden. Wenn eine Plausibilisierung der bestabgerechneten Effizienzwerte angestrebt wird, sollten Sensitivitätsanalysen durchgeführt werden, bei welchen die Effizienzwerte unterschiedlicher Modellkandidaten miteinander verglichen werden.²³ Hierbei ist es sinnvoll, beide Methoden, also DEA und SFA, getrennt voneinander zu validieren und nach den jeweiligen Modellspezifika optimal auszugestalten. Die DEA-Effizienzergebnisse müssen einer vertieften Peeranalyse unterzogen werden, welche die Outputgewichte sowie die Unternehmensgewichte (Lambdas) untersucht. Dabei gilt es sicherzustellen, dass die DEA-Effizienzwerte nicht durch einzelne Unternehmen maßgebend beeinflusst werden und/oder einzelne Strukturparameter für die Mehrheit der Unternehmen ohne Relevanz sind. So waren in der dritten Regulierungsperiode die Parameter Rohrvolumen und Jahreshöchstlast für die meisten VNB nur mit einem geringen Gewicht behaftet.

- Es sollten getrennte Sensitivitätsanalysen für DEA und SFA durchgeführt werden, bei welchen die Effizienzwerte unterschiedlicher Modellkandidaten miteinander verglichen werden.
- Bei der DEA muss sichergestellt werden, dass Effizienzwerte nicht durch einzelne Unternehmen maßgebend beeinflusst werden und/oder einzelne Strukturparameter für die Mehrheit der Unternehmen ohne Relevanz sind

²³ Analog zu einer „Efficiency Contribution Measure“ Analyse, vgl. Pastor, J., Ruiz, J. & Sirvent, I. (2002), A Statistical Test for Nested Radial Dea Models, *Operations Research*, Vol. 50(4): 728-735.