

[REDACTED]

Bundesnetzagentur
Beschlusskammer 8
Postfach 80 01
53105 Bonn

[REDACTED]

[REDACTED]

Datum: 25.01.2019

Per Mail

**Verwaltungsverfahren nach § 29 Abs. 1 EnWG i.V.m. § 4 Abs.1 und 2, § 32 Abs. 1 Nr. 1 und 5 ARegV
hier: Anhörung gemäß § 67 Abs. 1 EnWG**

Sehr geehrter Herr Bourwieg,
sehr geehrte Damen und Herren,

in Bezug auf die Veröffentlichung des Entwurfs des Gutachtens zum Effizienzvergleich Strom räumen Sie die Möglichkeit zur Stellungnahme ein, welche wir mit diesem Schreiben nutzen wollen:

Wir sehen die von Ihnen angekündigte Entscheidung in verschiedenen Punkten als nicht sachgerecht bzw. nicht zulässig an:

I: Datenqualität und -validierung

Um einen sachgerechten Effizienzvergleich durchzuführen, wird ein belastbarer Datensatz benötigt. Die BNetzA hat ihre Datenerhebung bis in den Herbst 2018 hinein plausibilisiert. Der im Juli 2018 erstmals veröffentlichte KTA-Datensatz wurde im Oktober 2018 nochmals veröffentlicht. Durch die Rücknahme der Schwärzungen wies er die Daten weiterer Netzbetreiber auf, wodurch eine nahezu vollständige Replikation der Analysen der Gutachter bzw. der BNetzA möglich ist. Aktuell hat die BNetzA den Datensatz nach einem Beschluss des Bundesgerichtshofs aufgrund einer Rechtsstreitigkeit um die Veröffentlichung von Daten nach § 31 ARegV von ihrer Website entfernt.



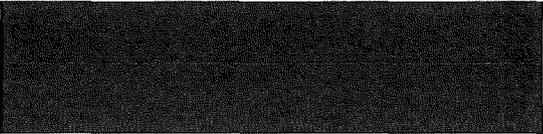
Die zeitlich jüngste Datenbasis stellt der am 5. Dezember 2018 von der BNetzA veröffentlichte Malmquist-Datensatz dar, welcher ebenfalls kurz darauf von der Internetseite der BNetzA entfernt wurde. Bei einzelnen Netzbetreibern sind darin Anpassungen von Kosten- und Strukturparametern gegenüber dem KTA-Datensatz zu beobachten, die wir auf weitere Plausibilisierungen und Kostenprüfungsabschlüsse zurückführen. Die Malmquist-Daten repräsentieren mit den Effizienzparametern der dritten Regulierungsperiode (RP3) aber nur einen sehr geringen Anteil der Positionen des KTA-Datensatzes, so dass sich die Beurteilung der Datenqualität sowohl auf den jüngsten KTA-Datensatz wie auch auf den Malmquist-Datensatz stützt.

1.1 Veröffentlichung, Aufbau und Transparenz der Datengrundlagen

Am 5. Oktober 2018 veröffentlichte die BNetzA eine erweiterte Version der am 24. Juli 2018 vorgestellten netzbetreiberindividuellen Aufwands- und Vergleichsparameter für den Effizienzvergleich der 3. Regulierungsperiode der Stromverteilernetzbetreiber. Die Datei umfasst nach wie vor 204 Stromverteilernetzbetreiber, wovon lediglich noch 2 Datensätze aufgrund laufender gerichtlicher Verfahren geschwärzt sind (vorher 26). Bezüglich aller weiteren Positionen ist die neue Version jedoch identisch zum Datensatz vom 24. Juli 2018. Für drei Netzbetreiber liegen nach wie vor noch keine Angaben zu den Gesamtkosten (und damit auch nicht zu standardisierten Gesamtkosten) vor. Es fehlen weiterhin Angaben zu insgesamt 70 Variablen in den Bereichen Einspeisung aus EEG/KWKG-Anlagen sowie Ausspeisung an Letztverbraucher. Im Gutachtenentwurf Effizienzvergleich Verteilernetzbetreiber Strom vom 21.12.2018 betonte das von der BNetzA mit der Durchführung des Effizienzvergleichs beauftragte Beraterkonsortium, dass die Aufwands- und Vergleichsparameter abschließend validiert seien (vgl. Swiss economics et al. Seite 54). Von den ursprünglich 204 Netzbetreibern fallen somit zwei aufgrund der Schwärzung und drei wegen fehlender Aufwandparameter heraus. Weiterhin wurde im Anschluss an die Kostentreiberanalyse ein Netzbetreiber ausgeschlossen. In Konsequenz dessen gehen 198 Netzbetreiber als Grundlage für den Effizienzvergleich in die Betrachtung ein (vgl. Swiss economics et al. Seite 54).

Der KTA-Datensatz 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.

Wie eingangs beschrieben, liegt aufgrund der Entscheidung des Bundesgerichtshofs derzeit noch kein finaler Stand des KTA-Datensatzes öffentlich vor. Der Vergleich der Kosten- und Strukturparameter aus der zuletzt veröffentlichten Malmquist-Datei mit dem KTA-Datensatz lässt darauf schließen, dass die Datenbasis noch bis Anfang Dezember plausibilisiert und vervollständigt wurde. So weisen 34 Netzbetreiber geänderte TOTEX und/oder sTOTEX auf, die Zählpunkte



haben sich bei 16 Netzbetreibern geändert, die Stromkreislängen MS und NS sowie die zeitgleichen Jahreshöchstlasten und die dezentralen Erzeugungsleistungen haben sich jeweils bei bis zu fünf Netzbetreibern verändert. Es ist somit zwar eine Anpassung und geringfügige Glättung des mutmaßlich verwendeten Datensatzes erkennbar – die im Sommer 2018 zur Konsultation festgestellten Strukturunterschiede bestehen aber weiterhin und werden daher in Abschnitt 1.3 und 1.4 detaillierter dargestellt.

Bei solchen nachträglichen Anpassungen sollten die Strukturdatenquittungen dem betroffenen Netzbetreiber aktualisiert und in vollständigem Umfang erneut bereitgestellt werden – einschließlich der gebietsstrukturellen Quittungsdaten.

Die Netzbetreiber 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. Eine solche Quittung, aus der (a) die konsolidierten Netzkosten und die dauerhaft nicht beeinflussbaren Kostenanteile, jeweils nach Kostenarten gegliedert, und die daraus abgeleiteten Aufwandsparemeter 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, liegt von den Landesregulierungsbehörden Bayern, Baden-Württemberg, Hessen, Nordrhein-Westfalen und Rheinland-Pfalz vor. Auch die BNetzA und die übrigen Regulierungsbehörden sollten den Netzbetreibern in ihrer Zuständigkeit diese Kostenquittungen nicht nur in Schriftform oder im PDF-Format, sondern auch als Excel-Datei 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.

Vor diesem Hintergrund sowie der im Folgenden dargestellten Auffälligkeiten ist aus Branchensicht die Veröffentlichung des finalen Standes zur endgültigen Beurteilung notwendig.

Forderungen:

Zur Erhöhung der Transparenz müssen die Netzbetreiber über die durchgeführten Anpassungen in elektronischer Form informiert werden. Strukturdatenquittungen sind nach Datensatzanpassungen dem betroffenen Netzbetreiber aktualisiert und vollständig auszuhändigen. Kostenquittungen sind den Netzbetreibern im Excel-Format

auszuhändigen. 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“.**

1.2 Behördliche Plausibilisierung

In der BNetzA-Unterlage „Konsultation zur Auswahl der Vergleichsparameter zum Effizienzvergleich der deutschen Elektrizitätsverteilernetzbetreiber“ vom 25. Juli 2018 wurde bereits angegeben, dass umfangreiche netzbetreiberindividuelle Plausibilisierungen vorgenommen wurden. Im Gutachtenentwurf Effizienzvergleich Verteilernetzbetreiber Strom vom 21. Dezember 2018 widmet sich Kapitel 4 dem Thema der Datenbearbeitung und -validierung.

Neben einer Vielzahl an formalen Prüfungen der Vollständigkeit, Datenformate und Wertebereiche wird vor allem Wert auf quantitative und analytische Prüfungen gelegt. Hierbei werden zum einen Logikprüfungen und Vergleiche mit bereits gelieferten, externen Daten aus anderen Verfahren und Altdaten vorgenommen, wobei nicht darauf eingegangen wird, was die Vergleichsbasis darstellt. Weiterhin werden zahlreiche Verhältnisse geprüft, sowohl hinsichtlich der Aufwandparameter als auch im Bereich der Vergleichsparameter. Außerdem werden die Stückkosten, d.h. Kombinationen aus Aufwands- und Vergleichsparametern sowie die Anzahl der Bestergebnisse bezogen auf die Stückkosten je Netzbetreiber betrachtet und auf Auffälligkeiten untersucht. Im Fazit wird festgehalten, dass die Daten formal korrekt, integer und plausibel zu sein scheinen (vgl. Swiss economics et al. Seite 54).

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



Laut BNetzA-Unterlage „Konsultation zur Auswahl der Vergleichsparameter zum Effizienzvergleich der deutschen Elektrizitätsverteilernetzbetreiber“ vom 25. Juli 2018 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 (vgl. Sumicsid et al. Seite 31) betroffen waren.

Forderungen:

Die Datenquellen der herangezogenen öffentlichen Daten sind zu nennen und offen zu legen.

Die Durchführung 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.

1.3 Analyse der Datengrundlage

Im Rahmen des BDEW/VKU/GEODE-Projekts „Benchmarking Transparenz“ (BMT) wurde eine gesonderte Untersuchung aller Parameter vorgenommen, die in den ursprünglichen Grundmodellen zur Anwendung kamen. Die Grundmodelle, die in der Konsultation mit „RP2“, „BU1“ und „TD2“ vorgestellt wurden, umfassten insgesamt 20 Strukturparameter:

Für diese Parameter wurden die Angaben der am BMT-Projekt teilnehmenden Unternehmen in der KTA-Datei auf mögliche Unregelmäßigkeiten untersucht. Einzelne Unregelmäßigkeiten wurden seit der Konsultation offenbar ausgeräumt.

Eine über diese Modellparameter hinausgehende deskriptive Analyse sämtlicher Parameter der KTA-Daten zeigt jedoch nach wie vor, 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

- Industrienetzbetreiber: Eine Versorgung der eigenen Produktion mit infrastrukturellen Dienstleistern ist stets mit einer sehr geringen Anzahl von Zählpunkten bei gleichzeitig hohen übertragenen Energiemengen verbunden (Verhältnis von Zählpunkten oder Bevölkerungszahl zu Jahreshöchstlast (JHL) gleich null; extreme Abweichungen bei JHL pro Anschlusspunkt (> 10.000 % im Median) im Vergleich zu anderen Verteilernetzbetreibern). Darüber hinaus betreiben diese nur eine oder zwei Spannungsebenen und es erfolgt – im Unterschied zu „regulären“ Stromverteilernetzbetreibern – keine Einspeisung aus dezentralen, regenerativen Energiequellen.
- Dezentrale installierte Leistung zu JHL: Schwankungen um den Mittelwert (95%) in einer Spanne von 0 % bis 400 %
- Einspeisepunkte: Netzbetreiber mit nur einem Einspeisepunkt von Erzeugungsanlagen weisen extreme Abweichungen im Verhältnis zur zeitgleichen JHL im Vergleich zu anderen Verteilernetzbetreibern aus
- Ausspeisungen NS zu Jahreshöchstlast Ausspeisungen NS: Schwankungen im Bereich -35 % bis +50 %
- NS-Freileitungen in m pro Mast/Dachständer: große Abweichungen mit 15 bis 60 und mehr Meter pro Mast
- MS-Freileitungen in m pro Masten: große Abweichungen mit 50 bis 200 und mehr Meter pro Mast
- Anteil NS HAL Kabel an NS Kabel: Schwankungen im Bereich -20 % bis +20 % zum Median
- Anteil NS HAL Freileitungen an NS Freileitungen: Werte im Bereich 0 % bis 90 %
- Stromkreislänge NS pro versorgter Fläche: Werte im Bereich 10 km bis 40 km pro versorgtem km²
- Anteil RONT an Trafoleistung MS/NS: ein 100. %iger Anteil der Regelbarkeit unter Last ist bei den Transformatoren sehr unwahrscheinlich
- Anzahl Anschlusspunkte pro Ortsnetzstation: Streuung der Werte von <5 bis 15
- Anteil Pauschalanlagen an Zählpunkten NS: einige wenige NB mit sehr großen Abweichungen zum Rest (nur drei NB größer 2 %, zwei NB mit knapp 15 %)

- 
-
- Anteil Anschlusspunkte Straßenbeleuchtung an Anschlusspunkten gesamt: Median 1,3 % und dabei einige NB > 25 %.

Forderungen:

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 gegenüber den Netzbetreibern dokumentieren.

Im Falle der Identifikation besonderer Umstände, die zu Auffälligkeiten bei den einzelnen Netzbetreibern führen, ist zu prüfen und zu dokumentieren, ob die grundsätzliche Vergleichbarkeit dieser Netzbetreiber gegeben ist oder ob ein Ausschluss jener Unternehmen bei der Ermittlung der Effizienzwerte angemessen wäre.

1.4 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 dienten das 1., 5., 95. und 99. Perzentil als Orientierung für die Verteilung der Werte über die Stichprobe hinweg. Eine Untersuchung der Netzbetreiber, die besonders häufig 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 Parameterwerte konzentrieren sich hier auf einige wenige Netzbetreiber, die in bis zu 352 Variablen zu den höchsten Werten zählen. Dies spiegelt die erheblichen Größenunterschiede am Markt wider und unterstreicht die Heterogenität der Größe von Netzbetreibern in der Stichprobe.

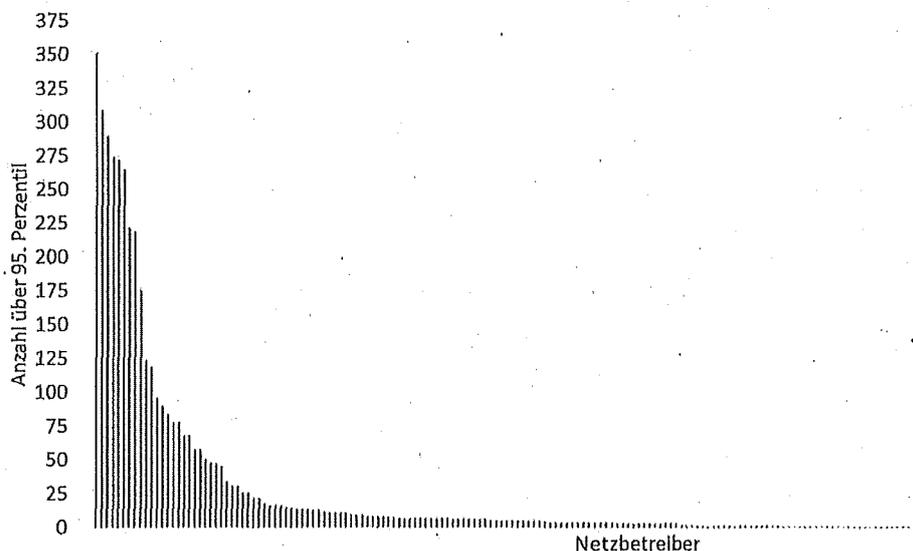


Abbildung 1: Anzahl der Überschreitungen des 95. Perzentils in den einzelnen Parameterwerten

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 den 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 Seite 54 des Gutachtenentwurfs, 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.

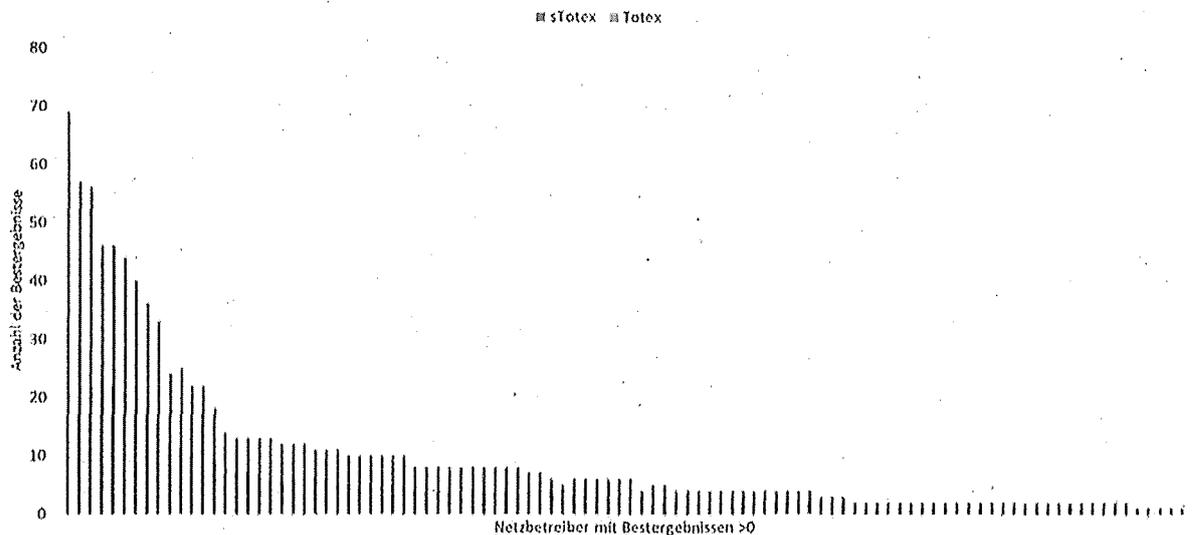


Abbildung 2: Netzbetreiber mit einem oder mehr Bestwerten im Ranking der Unit Costs, sortiert nach Anzahl der Bestwerte in sTOTEX-basierter Rechnung (exkl. geschwätzter Variablen und Variablen, die mit Ja/Nein beantwortet wurden)

Die 14 Netzbetreiber mit den häufigsten Bestergebnissen fallen besonders dadurch auf, dass diese entweder nur sehr wenige Spannungsebenen bedienen und insgesamt sehr kleine oder aber besonders große Netzbetreiber sind (vgl. dazu auch die Definition von „kleinen“ und „großen“ Netzbetreibern durch die BNetzA, Seite 53 des Gutachtenentwurfs). 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.

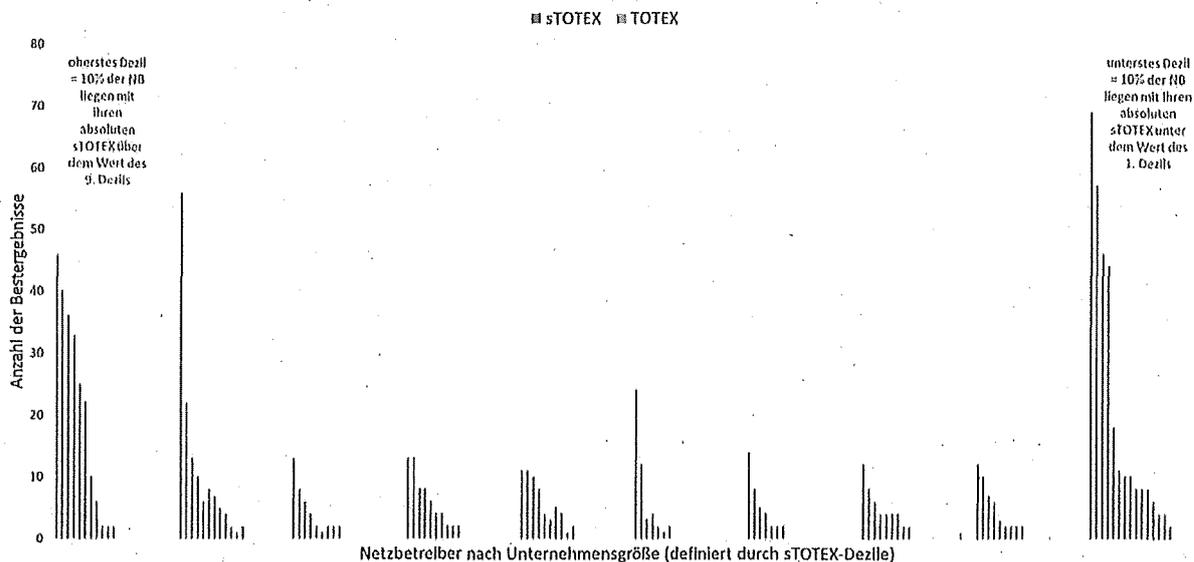


Abbildung 3: Netzbetreiber nach Bestwerten im Ranking der Unit Costs nach Anzahl der Bestwerte in sTOTEX-basierter Rechnung und Dezil der sTOTEX-Höhe (exkl. geschwätzter Variablen und Variablen, die mit Ja/Nein beantwortet wurden)

Wie auch im Gutachtenentwurf erläutert, wurden als Stromverteilernetzbetreiber mit besonders vielen Bestergebnissen vermehrt „kleine“ oder „große“ Netzbetreiber identifiziert. Dieses wird auch aus Abbildung 3 ersichtlich: bei besonders „großen“ 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. Swiss economics et al. Seite 38), wie die Verteilungen ganz rechts zeigen.

„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 bei der Ermittlung von sachgerechten Effizienzwerten für alle Stromverteilernetzbetreiber ggf. zu berücksichtigen ist.

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. Sumicsid 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 Bestsergebnisse 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.

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

Datenauffälligkeiten hinsichtlich struktureller Unterschiede sollten detailliert untersucht werden. Sollte die BNetzA zu dem Schluss kommen, dass Datenkorrekturen notwendig sind, sind diese den betroffenen Netzbetreibern mitzuteilen. Falls keine Korrekturen notwendig sind, ist mindestens im Gutachten aufzuzeigen, warum die gefundenen Auffälligkeiten plausibel sind und wie mit ihnen im weiteren Verlauf der Effizienzwertermittlung umgegangen wird.

II. Kostentreiberanalyse

2.1 Methodische Vorgehensweise

Das Beraterkonsortium der Bundesnetzagentur führt eine Vorauswahl von Vergleichsparametern aufgrund einer statistischen Kostentreiberanalyse durch, wobei unterschiedliche Verfahren angewendet werden (Bottom-up Analyse ohne Priorisierung der Daten, Top-down Analyse ausgehend von einer Priorisierung der möglichen Vergleichsparameter sowie eine Bewertung des Modells der zweiten Regulierungsperiode). Ausgehend hiervon werden einige Modelle zur „händischen“ Weiterentwicklung ausgewählt.

Die Branche begrüßt es, dass im Vergleich zu dem zur Konsultation vorgeschlagenen Vorgehen einige Anpassungen am Modellfindungsverfahren vorgenommen wurden. Beispielsweise wurden in der „Top Down“ und „Bottom Up“ Analyse unterschiedliche funktionale Formannahmen berücksichtigt und Pfadabhängigkeiten zumindest in der „Bottom Up“ Analyse durch die Anwendung von LASSO-Methoden verringert. De facto hat dies jedoch für die Effizienzmessung keine Bedeutung, da die Modelle der „Bottom Up“ und „Top Down“ Untersuchung letztendlich ohnehin nicht verwendet werden.



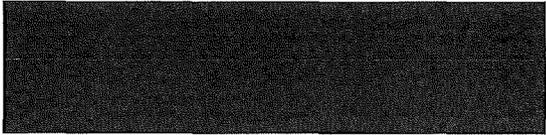
An dem Verfahren zur Findung des Modells ist jedoch weiterhin zu kritisieren, dass die Berater der Bundesnetzagentur – letztendlich mit dem Ziel, einseitig ein vermeintliches Risiko einer Überspezifizierung des Modells zu vermeiden – auf unnötig stringente und stellenweise nicht anwendbaren statistischen Methoden beharren: Beispielsweise sind die Überlegungen zu einer „optimalen“ Anzahl von Vergleichsparametern nach wie vor abzulehnen (vgl. Kapitel 2.3), es wird ein zu starker Fokus auf die Vermeidung von Multikollinearität gelegt (vgl. Kapitel 2.2) und die Methoden zur Überprüfung und Validierung der Ergebnisse (Second Stage Analysen, vgl. Kapitel 4.3) sind grundsätzlich nicht anwendbar. Viele der angewendeten Testverfahren wurden bereits in der Vergangenheit von der Branche stark kritisiert. Die Kritik an diesen Verfahren ist weiterhin uneingeschränkt gültig.

In Anbetracht der erheblichen Heterogenität der Netzbetreiber wird dem Risiko einer Unterspezifizierung des Modells (d.h. zu wenige Vergleichsparameter werden bei der Effizienzmessung berücksichtigt) im Gutachtenentwurf relativ wenig Bedeutung beigemessen.

Von einem ausgewogenen Bericht wäre zu erwarten gewesen, dass die Risiken beider Spezifizierungsfehler in Betracht gezogen werden. Grundsätzlich sollte bei der Modellwahl immer beachtet werden, dass in der wissenschaftlichen Literatur das Risiko eines unterspezifizierten Modells als deutlich relevanter angesehen wird als das Risiko eines überspezifizierten Modells (vgl. z.B. Smith, 1997, S. 250, Galagedera und Silvapulle, 2003, S. 657). Es ist daher zu erwarten, dass geeignetere Modelle zur Modellauswahl gefunden werden könnten, die bei der Überarbeitung des Gutachtens berücksichtigt werden sollten (vgl. Kapitel 4.2).

Auch die novellierte Version der ARegV fordert explizit die weitgehende Abbildung von Heterogenität (§ 13 Abs. 3 ARegV). Dabei erkennen die Berater die stark unterschiedliche Struktur der Netzbetreiber durchaus und stellen diese auch dar. So wird beispielsweise in Abbildung 12 (S. 49, Swiss Economics et al. 2018) dargestellt, wie viele Netzbetreiber jeweils auf welchen Netzebenen tätig sind. Dabei stellt sich heraus, dass beispielsweise nur wenige Netzbetreiber auch die Höchstspannung bedienen und einer ausschließlich die Hochspannung bewirtschaftet. Inwiefern diese Netzbetreiber überhaupt in ihrer Kostenstruktur mit der „Mehrheit“ vergleichbar sind, wäre zu untersuchen. Die Kostentreiberanalyse nimmt darauf jedoch keinen weiteren Bezug. Durch das letztlich gewählte Modell sind zwar alle Spannungsebenen vertreten (vornehmlich wegen des Kriteriums der „Vollständigkeit“), an keiner Stelle wird aber gezeigt, dass dies zur korrekten Abbildung der stark unterschiedlichen Struktur der Netzbetreiber ausreicht. Mehr noch wird auf Ebene der Mittelspannung eine kaum begründbare Aggregation vorgenommen, indem nur ein Parameter für Freileitungen und Kabel zum Einsatz kommt.

Nicht zuletzt scheint ein sehr hoch gewichtetes Kriterium ein signifikanter Ineffizienzterm in der SFA gewesen zu sein. Mehrere Entscheidungen hinsichtlich der Parameterwahl wurden offenbar unter diesem Aspekt getroffen, namentlich auch die noch weiter zu diskutierende Aggregation von Freileitungen und Kabeln in der Mittelspannung. Es ist zu beachten, dass nach zwei Perioden bzw. zehn Jahren Anreizregulierung mit individuellen Effizienzvorgaben durchaus die Möglichkeit



besteht, dass in der Branche die Ineffizienz bis hin zu einem nicht mehr signifikanten Maß abgebaut wurde. Wenn also energiewirtschaftlich und statistisch sinnvolle Modelle in mehreren Fällen eine solche signifikante Ineffizienz nicht mehr ausweisen, kann es nicht das Ziel sein, so lange Modelle zu verändern, bis wieder Ineffizienz vorliegt. Im Falle der Mittelspannungsebene kann beispielsweise durch die Aggregation des Parameters theoretisch Ineffizienz einfach dadurch erzeugt werden, dass Unternehmen mit einem hohen Verkabelungsgrad, verglichen mit Unternehmen mit einem hohen Anteil Freileitungen, ineffizient erscheinen. Wie in den anderen Netzebenen auch der Fall, gibt es bei den Mittelspannungsleitungen energiewirtschaftliche und exogene Gründe, die es erforderlich machen, situationsbedingt entweder Mittelspannungskabel zu verlegen oder Mittelspannungsfreileitungen zu bauen, während die damit verbundenen Kosten differieren.

2.2 Parameterwahl DEA/SFA und Multikollinearität

Die BNetzA-Berater betonen an verschiedenen Stellen, dass die Multikollinearität im Effizienzvergleich möglichst zu begrenzen sei (z. B. Seite 65). Dieser Ansatz ist im Kontext der Effizienzmessung jedoch einerseits ein Fehlschluss, andererseits führt er zu diversen nicht nachvollziehbaren Konsequenzen auch im Kontext der Effizienzmessung in der DEA. Wie die Verbände bereits in diversen Stellungnahmen erläutert haben, stellt Multikollinearität 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 alleinstehend, sondern nur in Kombination aller Koeffizienten der kollinearen Modellparameter interpretiert werden können. Multikollinearität hat 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).

Die Tatsache, dass Multikollinearität im Kontext der Effizienzmessung eine untergeordnete Rolle spielen muss, wird erstmalig auch von Seiten der Berater der BNetzA im Gas-Effizienzvergleich anerkannt (Frontier Economics und Technische Universität Berlin, 2018, S. 110: «Im Zusammenhang mit der Schätzung individueller Ineffizienz kommt dem Problem der Multikollinearität ohnehin eine geringe Bedeutung zu, da die Interpretation der Koeffizienten nur zweitrangig ist»).

Die Beibehaltung des Kriteriums der Multikollinearität hat einige problematische Konsequenzen auf die Ermittlung der Effizienzwerte mittels der Methode der DEA. So kann gerade die auf Seite 23 des Gutachtens getätigte Schlussfolgerung, dass die „Anwendung des Vorsichtsprinzips“ die



Verwendung derselben Vergleichsparameter für beide Methoden notwendig machen, nicht nachvollzogen werden. Einerseits bedürfte das Vorsichtsprinzip gerade einer Anwendung von Kostentreiberanalysen, welche die Vor- und Nachteile der beiden Methoden spezifisch berücksichtigen. Andererseits wird aufgrund der Fokussierung auf eine Modellwahl mittels OLS die DEA stark eingeschränkt: Besonders deutlich wird dies beim auf Seite 36 des Gutachtens beschriebenen Vorgehen. So werden größenunabhängige Vergleichsparameter mit der unterlagerten Stromkreislänge skaliert, damit sie sowohl in der DEA als auch in der SFA verwendet werden können. Da diese Skalierungsvariable jedoch mit bereits im Modell vorhandenen Vergleichsparametern hochkollinear ist, besteht die Gefahr, dass die skalierte Variable nicht signifikant in der OLS ist und so für die DEA ausgeschlossen wird. An anderer Stelle werden mit dem Verweis, dass die DEA nur «positive Vorzeichen für die verwendeten Kostentreiber zulässt» (S. 68), Restriktionen in der Modellfindung eingeführt, obwohl die Parameter in der DEA nur einzeln eine positive Kostenwirkung aufweisen müssen (und keinesfalls dafür eine multiple Regression unter Multikollinearität herangezogen werden kann).

Auch die Aussage der Berater, dass die ARegV keine unterschiedlichen Vergleichsparameter in den beiden Methoden zulassen würde, kann nicht nachvollzogen werden. Die ARegV schreibt gleiche Parameter explizit für beide Kostenbasen vor, nicht aber für die Methoden. Der Fokus auf die zwingend gleichen Parameter schwächt die eingesetzten Methoden, da in der SFA Parameter kategorisch ausgeschlossen werden, welche nicht für die DEA geeignet sind oder einer Skalierung bedürfen (s.o.) und damit eine zu hohe Korrelation mit anderen Parametern aufweisen. Demgegenüber wird in der DEA die mögliche Parameterzahl aufgrund von Kriterien, welche nur für die parametrische Methode zutreffen (bspw. Signifikanz der Parameter unter Multikollinearität), begrenzt.

Forderungen

Eine Auswahl von Vergleichsparametern ausschließlich anhand der statistischen Signifikanz steht im direkten Widerspruch zu der verstärkt definierten Anforderung aus der ARegV, die die Berücksichtigung der Heterogenität der Netzbetreiber zwingend vorschreibt.

Bei der Auswahl der Modelle müssen beide Risiken von Fehlspezifizierung (Über- und Unterspezifizierung) ausgewogen berücksichtigt werden. Der Umstand, dass relevante Vergleichsparameter im Modell noch nicht berücksichtigt wurden, muss bei der finalen Überarbeitung des Gutachtens berücksichtigt werden.

Für DEA und SFA müssen unterschiedliche Parameter zugelassen werden, um die Stärken der jeweiligen Methoden voll auszuspielen zu können.

2.3 Untersuchungen zur Modellgröße

Die BNetzA-Berater wenden das LASSO-Verfahren zur Indikation einer optimalen Modellgröße auf Basis einer linearen (vgl. Abbildung 20 auf S. 61) und einer logarithmierten (vgl. Abbildung 21 auf S. 61) Spezifikation der Variablen an. Fraglich ist, warum die Berater das LASSO-Verfahren nicht auf Basis der finalen funktionalen Form, d. h. insbesondere mit Zählpunkten normierten Variablen durchgeführt haben. Eine entsprechende Analyse anhand des aus dem Gutachten erschlossenen Vorgehens der BNetzA deutet auf eine wesentlich höhere Anzahl an optimalen Parametern hin. In Abbildung 4 ist die optimale Anzahl an Parametern auf Basis von mit Zählpunkten normierten Variablen dargestellt. Die Abbildung deutet auf eine optimale Anzahl an Parametern zwischen 13 und 29 Variablen hin (rot markiert) und lässt die im Gutachten ausgewiesene optimale Modellgröße von ca. 6 Parametern unter Berücksichtigung der abschließend gewählten funktionalen Form als fraglich erscheinen. Es erscheint also durchaus notwendig, auch Modelle mit einer höheren Anzahl an Parametern in Betracht zu ziehen.

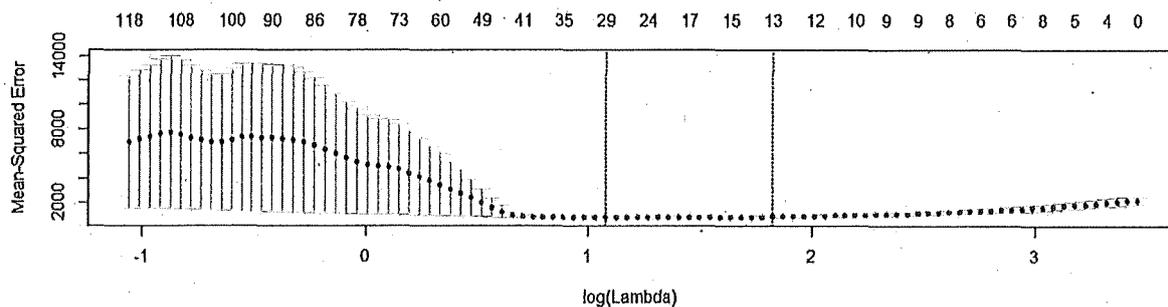


Abbildung 4: Optimale Anzahl Parameter – Auswertung BU-Modelle (linear, normiert mit Zählpunkten)

Forderung

Es müssen Modelle getestet werden, welche deutlich mehr Parameter als das bisher gewählte Modell enthalten. Eine umfassende Analyse mit den Methoden der Berater selbst weist deutlich in diese Richtung.

III. Durchführung der Ausreißeranalyse

Der Identifikation von Ausreißern im Rahmen einer Effizienzanalyse kommt große Bedeutung zu. Dieser Tatsache ist sich auch der Verordnungsgeber bewusst und spezifiziert in Anlage 3 zu § 12 Abs. 5 ARegV die Art und Weise, wie Ausreißer zu identifizieren sind. Die Bereinigung des Datensatzes um Ausreißer vor der Durchführung eines Effizienzvergleichs leitet sich aber auch aus § 21 a Abs. 2 Satz 4 EnWG ab, indem eine vergleichbare Datenbasis von objektiv strukturell vergleichbaren Unternehmen gefordert wird. Die Berater der BNetzA gehen in ihrem Gutachten (Swiss Economics et al. 2018) auf den Seiten 76 und 77 für die DEA sowie den Seiten 79 bis 81 für die SFA auf die vorgenommene Ausreißeranalyse ein. Grundsätzlich orientieren sich die Berater dabei an der Vorgehensweise in den bisherigen Effizienzvergleichen und verwerfen die von der Branche aufgebrachten Änderungsvorschläge. Die dabei vorgebrachten Argumente gegen die Änderungsvorschläge verfangen jedoch nicht. Im Folgenden fassen wir deshalb unsere Kritik an der bisherigen Vorgehensweise noch einmal zusammen und zeigen, dass das von den Beratern gewählte Vorgehen zu einer verzerrten und zu tiefen Durchschnittseffizienz führt.

3.1 Notwendige Anpassungen bei der Ausreißeranalyse

Die Identifikation von Ausreißern in der DEA erfolgt auf zwei Stufen (Anlage 3 zu § 12 Abs. 5 ARegV):

- Dominanzanalyse
- Supereffizienzanalyse

3.1.1 Dominanzanalyse

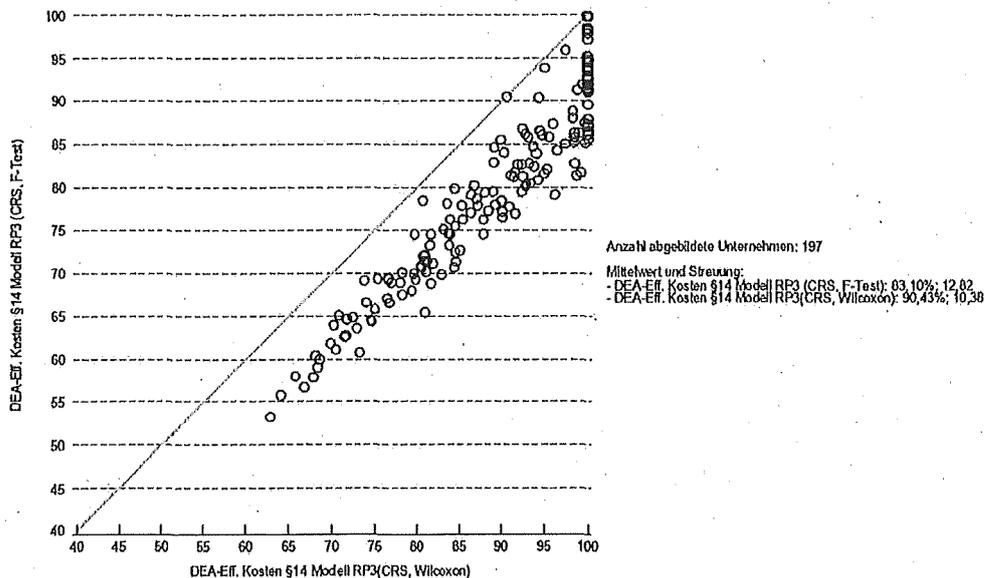
Bei der Dominanzanalyse geht es darum, Netzbetreiber zu identifizieren, die für einen überwiegenden Teil des Datensatzes als Effizienzmaßstab gelten. Der statistisch signifikante Unterschied in der mittleren Effizienz mit und ohne ein Unternehmen ist dabei mit einer Vertrauenswahrscheinlichkeit von 95 Prozent zu identifizieren. Die Gutachter der BNetzA beurteilen den Einfluss auf die mittlere Effizienz auf Basis des sogenannten Banker-Tests. Dieser Test wiederum basiert auf einer F-Verteilung. Diese Vorgehensweise geht auf Banker (1993) zurück. Er entwickelte einen Hypothesentest für «two groups of decision making units (DMU) to assess whether one group is more efficient than the other» (Banker, 1993, S. 1271). Dieser Test setzt voraus, dass die Effizienzwerte der beiden zu vergleichenden Gruppen unabhängig voneinander sind und die Verteilung der Effizienzwerte bekannt ist. Beides ist hier nicht der Fall. Insbesondere sind die Effizienzwerte der beiden Gruppen nicht unabhängig, da die Unternehmen



in beiden Gruppen identisch sind. Zusätzlich besteht bei der Anwendung des Banker-Tests ein Problem, falls die Daten nicht unerhebliche Messfehler enthalten. In diesem Fall ist der Banker-Test selbst dann nicht ideal, wenn dessen sonstige Voraussetzungen erfüllt wären. Banker schreibt: «We find that tests developed in this paper perform better than the tests in Banker (1993) when noise plays a significant role in the data generating process.» (Banker et al. 2010, S. 238). Bei den vorliegenden Daten der deutschen Netzbetreiber muss davon ausgegangen werden, dass Messfehler in einem gewissen Ausmaß vorliegen. Davon geht offenbar auch der Gesetzgeber aus, da er die Anwendung der SFA als stochastische Methode, deren Annahmen explizit Datenfehler voraussetzen, ausdrücklich vorgeschrieben hat. Um das Problem der Abhängigkeit der beiden Gruppen zu beseitigen, könnte anstelle des Banker-Tests der Wilcoxon signed-rank test angewendet werden. Dieser behebt das Problem der Berücksichtigung abhängiger Stichproben, allerdings wird auch hier unzutreffend von einer symmetrischen Verteilung der Differenzen der Ineffizienz ausgegangen. Bei durchgeführten ersten Tests erwies sich dieses Problem aber als nachrangig. Soll auch korrekterweise eine einseitige Verteilungsannahme unterstellt werden, böten sich beispielsweise Bootstrapping-Methoden an. Aufgrund dieser offensichtlichen Mängel haben die Verbände gefordert, dass die Dominanzanalyse auf Basis nicht-parametrischer Tests durchgeführt wird, welche die „paired“-Struktur der vorliegenden Daten berücksichtigen (z. B. Wilcoxon signed-rank test, SignTests, Bootstrapping-Methoden o. ä.).

Die Gutachter gehen nicht auf diese Kritik ein, sondern erwähnen lediglich, dass mögliche alternative Ausgestaltungen des Dominanztests diskutiert und verworfen wurden. Eine Begründung für dieses Vorgehen fehlt vollständig. Das Fehlen einer vertieften Diskussion der alternativen Ansätze ist umso erstaunlicher, als dass die Wahl des Dominanztests starke Auswirkungen auf die Anzahl der identifizierten dominanten Unternehmen und die ermittelten Effizienzwerte hat (vgl. Kapitel 3.2). Die Verbände fordern deshalb weiterhin eine Anpassung der gewählten Teststatistik. Dies nicht zuletzt auch deshalb, weil durch die Anwendung der falschen Teststatistik eigentlich dominante Unternehmen für viele Unternehmen den effizienzsetzenden Maßstab, d. h. den sogenannten Peer darstellen (vgl. hierzu Kapitel 3.2 zum Zusammenhang zwischen Dominanzanalyse und Peeranalyse).

Die Anwendung des Wilcoxon signed-rank test führt auf der Ebene der bestabgerechneten Effizienzwerte einerseits zu einer Erhöhung der Durchschnittseffizienz. Andererseits finden sich eine Vielzahl an Unternehmen, deren bestabgerechneter Effizienzwert sich zwischen fünf und zehn Prozentpunkte erhöht.



© 2018 Polynomix AG

Abbildung 5: DEA-Effizienzwerte mit F-Test und Wilcoxon signed-rank test in der Dominanzanalyse

Forderung

Bei der Dominanzanalyse ist eine wissenschaftlich adäquate Teststatistik anzuwenden. Damit können die wirklich dominanten Ausreißer identifiziert werden und die Effizienzgrenze wird nicht von wenigen (dominanten) Unternehmen definiert.

3.1.2 Supereffizienzanalyse

Nach der Durchführung der Dominanzanalyse und dem Ausschluss der dominanten Unternehmen wird von den Gutachtern, wie in der ARegV gefordert, eine Supereffizienzanalyse durchgeführt. Dabei gilt ein Netzbetreiber als Ausreißer, wenn dessen Effizienzwert den oberen Quartilswert um mehr als den 1,5-fachen Quartilsabstand übersteigt. In der Stellungnahme der Verbände vom August 2018 wurde aufgezeigt, dass das bisherige Verfahren zur Supereffizienzanalyse deutliche Mängel aufgrund der Problematik von „maskierten“ Ausreißern zeigt. In der wissenschaftlichen Diskussion wird dieses Phänomen unter den Begriffen „masking effects“ (ein Ausreißer wird nur ohne zweiten Ausreißer, der den ersten ansonsten „maskiert“, identifiziert) und „swamping effects“ (ein Ausreißer wird nur zusammen mit zweitem Ausreißer identifiziert) diskutiert (vgl. z. B. Acuna, 2004). Bei den gemäß ARegV aufgeführten Methoden für die Identifikation von Ausreißern in der DEA besteht die Gefahr, dass solche verdeckten Ausreißer nicht adäquat erkannt werden können,

was sich verzerrend auf die Effizienzermittlung auswirken kann. Die Existenz von „maskierten“ Ausreißern wurde auch vom BGH (Bundesgerichtshof, 2018, 40 ff.) erkannt.

In der wissenschaftlichen Literatur existieren diverse Möglichkeiten, um verdeckte Ausreißer zu identifizieren. Dazu gehören unter anderem zweistufige Vorgehen in der DEA (vgl. Hammerschmidt et al, 2009), Datenwolken-Methoden (vgl. Bogetoft et al., 2010) und adjustierte Boxplots (vgl. Hubert, 2008). Die Verbände teilen die Ansicht der Gutachter nicht, dass ein iteratives Vorgehen bei der Ausreißeranalyse nicht durch die ARegV gedeckt sei, ein solches Vorgehen kann also durchaus in Betracht gezogen werden. Auch schließt das BGH-Urteil zur Berücksichtigung der ehemaligen regionalen Fernleitungsnetzbetreiber dies nicht aus, sondern sieht vielmehr Anlage 3 der ARegV als Mindestanforderung. Entgegen der Aussage der BNetzA (Bescheide, 2 Regulierungsperiode) kommt es auch nicht zu einem endlosen „Abschälen der Zwiebel“, denn die Supereffizienzwerte aller Unternehmen normalisieren sich, wenn sämtliche Ausreißer aus der Berechnung entfernt werden. Dies zeigt sich auch bei den vorliegenden Daten:

Nach der einmaligen Anwendung der Supereffizienzanalyse werden sieben Unternehmen im Modell TOTEX und sechs Unternehmen im Modell sTOTEX als Ausreißer identifiziert und ausgeschlossen. Wiederholt man nach Ausschluss dieser Ausreißer die Supereffizienzanalyse erneut, werden bei TOTEX und sTOTEX vier weitere Unternehmen als Ausreißer identifiziert. Für die sTOTEX findet sich schließlich in der dritten Runde noch ein supereffizientes Unternehmen (vgl. Tabelle 1).

Supereffizienzanalyse	Ausreißer TOTEX	Ausreißer sTOTEX
Erstmalig	7	6
Zweite Runde	4	4
Dritte Runde	0	1

Tabelle 1: Erkannte Supereffizienzausreißer bei mehrfacher Durchführung der Supereffizienzanalyse

Wie die Tabelle 2 zeigt, resultieren für die Unternehmen, welche in der zweiten Runde bei den TOTEX als supereffiziente Unternehmen identifiziert werden, immer noch Supereffizienzwerte von über 200%. Das bedeutet, dass ein Unternehmen als Peer in Erscheinung tritt, welches auch nach Verdoppelung seiner Kosten immer noch einen Effizienzwert von 100 % erhalten würde und weiterhin Peer wäre.

Unternehmen	Supereffizienzwert
1	226 %
2	207 %
3	129 %
4	127 %

Tabelle 2: Supereffizienzwerte für die in der zweiten Runde identifizierten SE-Unternehmen mit Totex

Der Ausschluss dieser maskierten Ausreißer führt zu deutlichen Veränderungen der DEA-Effizienzwerte (Abbildung 6, beispielhaft mit TOTEX). Auch der bestabgerechnete Effizienzwert

erhöht sich im Durchschnitt, wobei sich für eine Vielzahl an Unternehmen ihr bestabgerechneter Effizienzwert zwischen fünf und zehn Prozentpunkte erhöht. Aufgrund der dargestellten Sensitivität der Effizienzwerte auf den Ausschluss zusätzlicher Ausreißer ist es nicht nachvollziehbar, wie die BNetzA im Gutachten zum Schluss kommt, dass „es vorliegend nur sehr geringfügige Auswirkungen hätte“ (Swiss Economics et al. 2018, s. 77).

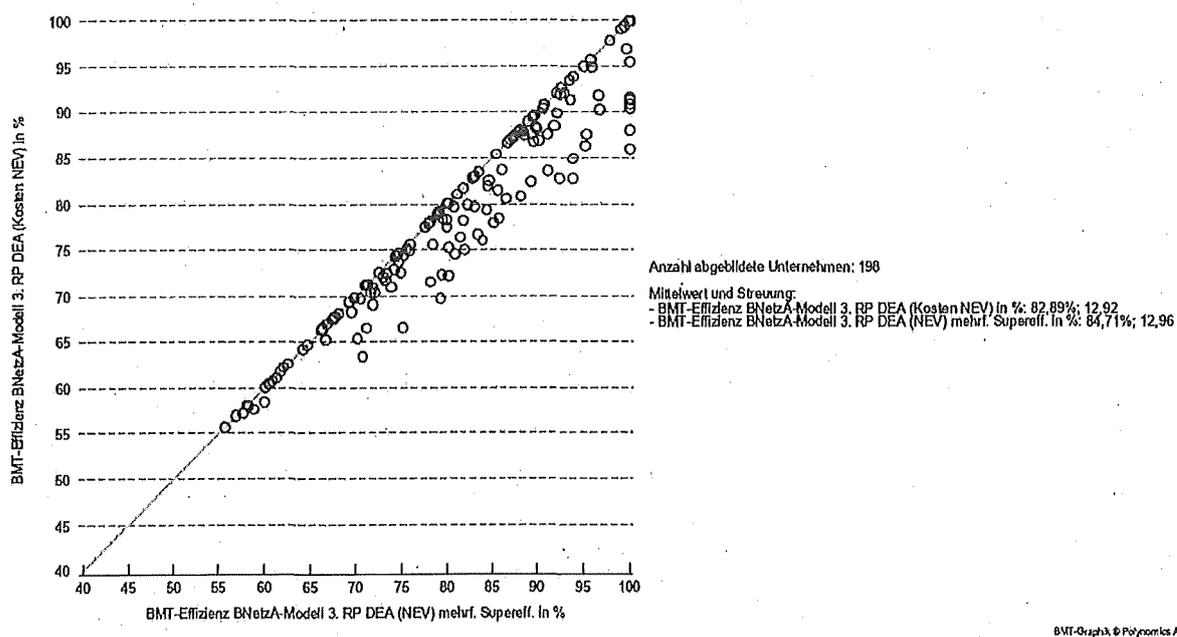


Abbildung 6: DEA-Effizienzwerte (TOTEX) mit einfacher versus mehrfacher Supereffizienzanalyse

Forderungen

Das Problem verdeckter Ausreißer („masking effects“ und „swamping effects“) ist in der Literatur bekannt und vom BGH anerkannt. Die Gutachter tragen dieser Problematik nach wie vor keine Rechnung. Die Ausreißeranalyse ist anzupassen, damit multiple Ausreißer entdeckt werden können. Die Ausführungen in Anlage 3 ARegV sind gemäß BGH-Urteil als Mindestanforderungen in Bezug auf die Ausreißeranalyse anzusehen.

3.2 Zusammenhang zwischen Dominanz- und Peeranalyse

Die Gutachter beschreiben auf Seite 100 sehr kurz die Ergebnisse ihrer Peeranalyse für die DEA-Berechnungen. Dabei zeigt sich, dass unter Ausschluss der Ausreißer einzelne Unternehmen für eine Vielzahl an Unternehmen als Peer im Sinne der „Best Practice“ gelten. So verdeutlicht Tabelle 27 des Gutachtens (S. 101), dass beispielsweise bei der DEA mit TOTEX ein Unternehmen (teils zusammen mit anderen Peers) für 88 oder ein weiteres Unternehmen für 70 Unternehmen den Peer darstellt. Von den 35 Peer-Unternehmen bei den TOTEX sind 11 Unternehmen für mindestens 20 Unternehmen Peer und von 37 Peer-Unternehmen bei den sTOTEX sind 10 Unternehmen für mindestens 20 Unternehmen die Peers. Die Gutachter verweisen darauf, dass diese Zahlen sich im Rahmen früherer Effizienzvergleiche bewegt.

Die teilweise hohe Anzahl an Verfolgern für einzelne Peers ist sehr auffällig und hätte seitens der Gutachter vertieft analysiert werden müssen. Bei den fünf Unternehmen, die jeweils bei mindestens einer Kostenbasis mehr als 50 Verfolger aufweisen, können die folgenden Auffälligkeiten identifiziert werden:

- Zwei Unternehmen weisen Anschlusspunkte nur für die MS und darunter aus, betreiben also praktisch nur diese Spannungsebenen. Keines der fünf Unternehmen betreibt die HöS/HS oder darüber.
- Vier der Unternehmen betreiben eigene Umspannstationen in nennenswerter Zahl nur in der MS/NS.
- Drei Unternehmen weisen insgesamt nur Netzlängen (über alle Ebenen) von teils deutlich unter 1000 km auf (Mittelwert im Datensatz deutlich über 6000 km). Die Mehrheit der Leitungen liegt dabei häufig auch in der NS.
- Drei Unternehmen weisen auch nur unter 50.000 Zähler aus (Mittelwert im Datensatz über 230.000).
- Die Netzkosten liegen bei drei Unternehmen auch deutlich unter 10. Mio. Euro (Mittelwert ca. 53 Mio. Euro).

Scheinbar führt hier also die mangelnde Berücksichtigung von Heterogenität dazu, dass Unternehmen mit einer gewissen Struktur für eine Vielzahl von Dimensionen (in diesem Fall für Parameter auf niederen Spannungsebenen und für solche, die nicht nach Spannungsebenen oder eben Betriebsmitteln differenziert werden) den Bestwert setzen. Dies erscheint auch logisch, da diese Unternehmen kostenseitig so günstig sind, weil sie gewisse Aufgaben, welche beispielsweise Netzbetreiber aufweisen, die über alle Spannungsebenen operieren, eben nicht zu bewältigen haben. Die Ausreißerkorrektur greift also hier zu kurz. Wie Kapitel 3.1.1 zeigt, vermag gerade die



Anwendung des Wilcoxon signed-rank test genau diese offensichtlich auffälligen Unternehmen zu identifizieren.

Hier zeigt sich der Zusammenhang zwischen der Dominanz- und der Peeranalyse. Wie in Kapitel 3.1.1 gezeigt, führt die Anwendung einer wissenschaftlich falschen Teststatistik dazu, dass vor der Durchführung des Effizienzvergleichs zu wenige dominante Unternehmen aus der Untersuchung ausgeschlossen werden. Diese dominanten und fälschlicherweise im Datensatz enthaltenen Unternehmen verzerren den Effizienzvergleich. Dies lässt sich an der Peeranalyse eindrücklich illustrieren.

In Tabelle 3 sind die Peer-Unternehmen mit der Anzahl ihrer Verfolger für die DEA mit TOTEX und für die DEA mit sTOTEX abgetragen. Bei dieser Peeranalyse wurde der von den Gutachtern verwendete F-Test berücksichtigt. Dabei zeigt sich, dass auch auf dem für die Berechnung zur Verfügung stehenden Datensatz eine ähnliche Verteilung der Peerunternehmen mit einer ähnlichen Verteilung der Anzahl Verfolger je Peer resultiert, wie sie die Gutachter auf ihrem Datensatz erhalten (vgl. Tabelle 27, S. 101 im Gutachten).

In den Spalten drei und sechs findet sich zudem die Information, welche dieser Peer-Unternehmen bei Anwendung der richtigen Teststatistik im Rahmen der Dominanzanalyse als Dominanzausreißer identifiziert werden. In den meisten Fällen handelt es sich dabei gerade um die Unternehmen, welche für eine Vielzahl an Unternehmen als Peer im Sinne der „Best Practice“ den Effizienzwert mitdefinieren.

TOTEX			STOTEX		
ID	Anzahl Verfolger	Dominanzausreißer Wilcoxon	ID	Anzahl Verfolger	Dominanzausreißer Wilcoxon
Ug. 101	89	ja	Ug. 201	79	ja
Ug. 102	70		Ug. 202	70	ja
Ug. 103	57	ja	Ug. 203	57	ja
Ug. 104	53	ja	Ug. 204	53	
Ug. 105	46	ja	Ug. 205	52	ja
Ug. 106	40	ja	Ug. 206	41	ja
Ug. 107	39	ja	Ug. 207	37	ja
Ug. 108	35	ja	Ug. 208	36	ja
Ug. 109	32	ja	Ug. 209	31	ja
Ug. 110	27	ja	Ug. 210	26	
Ug. 111	22		Ug. 211	18	ja
Ug. 112	13		Ug. 212	17	ja
Ug. 113	12		Ug. 213	15	ja
Ug. 114	12		Ug. 214	12	ja
Ug. 115	11		Ug. 215	12	
Ug. 116	10		Ug. 216	11	ja
Ug. 117	10		Ug. 217	11	ja
Ug. 118	9		Ug. 218	10	
Ug. 119	9		Ug. 219	10	ja
Ug. 120	8		Ug. 220	9	
Ug. 121	8		Ug. 221	9	
Ug. 122	7	ja	Ug. 222	9	ja
Ug. 123	6		Ug. 223	8	ja
Ug. 124	5	ja	Ug. 224	7	
Ug. 125	4		Ug. 225	7	
Ug. 126	3		Ug. 226	7	
Ug. 127	3		Ug. 227	4	
Ug. 128	3		Ug. 228	3	
Ug. 129	3		Ug. 229	2	
Ug. 130	2		Ug. 230	2	
Ug. 131	2		Ug. 231	2	
Ug. 132	1		Ug. 232	1	
Ug. 133	1		Ug. 233	1	
Ug. 134	1		Ug. 234	1	
Ug. 135	1		Ug. 235	1	
			Ug. 236	1	
			Ug. 237	1	

Tabelle 3: Peerunternehmen mit der Anzahl Verfolger bei Anwendung des F- und Wilcoxon signed-rank test in der Dominanzanalyse

Forderungen

Die Peeraanalyse zeigt deutlich, dass Unternehmen, welche sich strukturell spürbar vom Großteil des Datensatzes unterscheiden, für die Mehrheit der Unternehmen die Effizienzgrenze setzen. Dies kann als mangelnde Berücksichtigung der Heterogenität im Datensatz interpretiert werden. Auch hierdurch ergibt sich die Anwendung weiterführender Ausreißertests als zwingend.

IV. Beurteilung des BNetzA-Modells

4.1 Fehlende Transparenz bezüglich Sensitivitätsrechnungen

Basierend auf der Vorauswahl aus der Kostentreiberanalyse wird das Modell der zweiten Regulierungsperiode und der Top-down Methode händisch weiterentwickelt. Zur Effizienzmessung wird dann eine Weiterentwicklung des Modells der zweiten Regulierungsperiode verwendet. Die exakten Ergebnisse dieser einzelnen Berechnungsschritte (beispielsweise Effizienzwerte, Schätzkoeffizienten der SFA oder Testergebnisse, Tabellen 16 bis 21) werden jedoch nicht wiedergegeben.

Zudem verweisen die Gutachter auf Seiten 93 und 94 auf diverse Detailspezifikationen und Sensitivitätsrechnungen, die durchgeführt wurden. Ziel dieser Arbeiten war es gemäß Aussagen der Gutachter, sicherzustellen, dass kleine Änderungen des gewählten Modells „zu keinen Verbesserungen“ führen. Grundsätzlich ist es begrüßenswert, dass die Gutachter verschiedene Parameter- und Umsetzungsthemen testen. Die Ausführungen sind aber in verschiedener Hinsicht mangelhaft:

- **Keine Nennung der Beurteilungskriterien:** Die Gutachter verweisen lediglich darauf, dass keine Verbesserungen des gewählten Modells resultieren sollten. Was sie genau unter Verbesserungen verstehen, wird nicht dokumentiert. So ist unklar, ob damit eine Verbesserung der verschiedenen statistischen Gütemaße (einzelne oder alle), eine Erhöhung der Durchschnittseffizienz oder eine geringere Streuung der individuellen Effizienzwerte als Kriterium betrachtet wurde.
- **Keine konkrete Beschreibung der Vorgehensweise:** Auf Seite 93 und 94 werden zum einen aufgezählt, welche Parameterergänzungen und/oder -variationen am gewählten Modell vorgenommen wurden respektive welche Änderungen bei der Modellumsetzung geprüft wurden. Diese Aufzählung ist aber sehr allgemein gehalten. So ist oft unklar, welche Varianten beispielsweise der Leerstandskorrekturen getestet wurden oder ob



beispielsweise die Umspannebene HöS/HS zusätzlich oder als Ersatzparameter (und wenn ja für welchen) im gewählten Modell integriert wurde.

- **Keine transparente Darstellung der Ergebnisse:** Die Erkenntnisse aus den durchgeführten Sensitivitätsrechnungen werden nicht transparent dargestellt. Die im Gutachten auf S. 94 getroffene Aussage «gestützt auf die vorgenommenen Untersuchungen konnten die zur Anwendung kommende Ausreißeranalyse bestätigt werden, ebenso das in Abschnitt 5.5.3 gewählte Effizienzvergleichsmodell» ist ungenügend und nicht überprüfbar.

Forderung

Um das Vorgehen, die Ergebnisse und die Schlussfolgerungen aus den Sensitivitätsanalysen nachvollziehen zu können, müssten das Vorgehen, die Resultate und Beurteilungskriterien zumindest in einem Anhang transparent dargelegt werden.

4.2 Sensitivität der Ergebnisse auf die Parameterwahl

Nach Durchführung der Kostentreiberanalyse resultierte ein von den Gutachtern gewähltes Modell. Dieses wurde anschließend noch im Hinblick auf ausgewählte Vergleichsparameter und Modellumsetzungen getestet. Im Fokus standen dabei die folgenden Aspekte:

- **Zeitgleiche Höchstlast:** Es wurden Varianten gerechnet mit Leerstandskorrekturen und ergänzt um Rückspeisungen, dazu mit etwaiger Berücksichtigung der Umspannebene HöS/HS.
- **Abbildung dezentraler Erzeugung:** Es wurde unter anderem der Einfluss von alternativen Disaggregationen der Netzebenen, die Aufteilung der Einspeisepunkte statt der Erzeugerleistung oder verschiedene Ergänzungen mit installierter Erzeugerleistung und/oder Jahresarbeit geprüft.
- **Ausgewählte Ergänzungen:** Des Weiteren wurden verschiedene ausgewählte Ergänzungen wie die Berücksichtigung von Einspeisepunkten und/oder der Jahresarbeit sowie Dichteparameter getestet.
- **Modellumsetzung:** Bezüglich der Modellumsetzung standen das Testen unterschiedlicher funktionaler Formen und Normierungsvarianten, eine iterative Supereffizienzanalyse, die

Anwendung alternativer Prüfkriterien bei der SFA-Ausreißeranalyse oder der Einfluss von Netzbetreibern mit tiefen Effizienzwerten auf die Effizienzwertermittlung im Vordergrund.

Die Ergebnisse dieser Sensitivitätsanalysen werden weder präsentiert noch detailliert erläutert. Auch ist unklar, warum gewisse Sensitivitäten getestet wurden und andere nicht. Einige der erwähnten Punkte scheinen nachrangig oder aus ingenieurwissenschaftlicher Sicht irrelevant, andererseits sind weitere Tests über die genannten Parameter hinaus denkbar und sinnvoll. Es ist daher nicht nachvollziehbar, warum die Berater der Bundesnetzagentur davon ausgehen, dass das gewählte Modell die erhebliche Heterogenität der Stromverteilnetzbetreiber bestmöglich abbildet und dass Modellanpassungen zu keiner „signifikanten Verbesserung“ gegenüber dem gewählten Modell führen. Um die Auswahl der Modelle nachvollziehen und die Robustheit der Ergebnisse beurteilen zu können, sollten die Ergebnisse im Haupttext oder zumindest im Anhang des Dokuments wiedergegeben und ausführlich analysiert werden.

Beim Nachvollzug der im Gutachten erwähnten Sensitivitätsrechnungen und beim Test weiterer Alternativmodelle können unserer Ansicht nach durchaus Verbesserungen der einen oder anderen Art (durchaus auch in Bezug auf die statische Beurteilung) gegenüber dem von den Gutachtern gewählten Modell gefunden werden. Einige Parameter werden von den Gutachtern selbst im Rahmen der Kostentreiberanalyse als *interessant eingestuft, eine weiterführende Berücksichtigung* in der Kostentreiberanalyse ist jedoch nicht dokumentiert. Auch wenn die hohe Gewichtung der Kontinuität (und damit die letztlich Wahl eines Modells auf Basis der Parameter des Modells der zweiten Periode) zu begrüßen ist, hätten interessante Parameter, welche beispielsweise bei der Bottom-Up-Analyse (und der Anwendung der LASSO-Methode) gefunden wurden, einer weitergehenden Prüfung bedurft.

Die Verwendung diverser, unterschiedlicher Parameter führt in unterschiedlichem Maße zu Modellverbesserungen und zwar dahingehend, dass sich die ermittelten Effizienzwerte verbessern und die statistischen Gütemaße insgesamt eine bessere oder zumindest ähnliche Eignung anzeigen. Zusammenfassend lassen sich zwei Schlussfolgerungen ziehen:

1. Dass die Aussage der Berater, dass die getesteten Sensitivitäten in keinem Fall zu Verbesserungen geführt hätten, nicht nachvollzogen werden kann.
2. Dass eine sachgerechte und weitergehende Auswahl an Sensitivitätstests und die korrekte und vollständige Durchführung dieser, weitere Modellvarianten offenlegt, die sehr wohl zu Verbesserungen führen können.

Wie in Kapitel 2.2 gezeigt, ist die statistische Signifikanz einzelner Parameter in diesem Zusammenhang gegenüber der adäquaten Berücksichtigung von Heterogenität ein deutlich nachrangiges Beurteilungskriterium. Auch die gewählte Parameterbegrenzung ist nicht zwingend,



da bei LASSO-Untersuchungen mit einer normierten Funktion weit mehr Parameter als optimal ermittelt werden.

Forderungen

Es gibt energiewirtschaftlich und statistisch sinnvolle Modellalternativen zum gewählten BNetzA-Modell. Es ist durch die BNetzA aufzuzeigen, dass das letztlich gewählte Modell zu keiner Benachteiligung von Netzbetreibern führt.

Es ist generell aufzuzeigen, dass die Heterogenität im Datensatz hinreichend berücksichtigt wurde und nicht Unternehmen mit bestimmten Versorgungsaufgaben benachteiligt werden.

4.3 Beurteilung der Second Stage Analyse

Auch im Rahmen des dritten Effizienzvergleichs für die Stromverteilnetzbetreiber wurden die Ergebnisse mit Hilfe von Second-Stage-Analysen plausibilisiert (Swiss Economics et al, S. 102ff). Wie bereits in der Stellungnahme der Verbände vom August 2018 erläutert, geht aus der wissenschaftlichen Literatur hervor, dass die Second-Stage-Analyse nicht zur Modellvalidierung oder -plausibilisierung entwickelt wurde.

Dies vorangestellt ist zu erwähnen, dass die in Tabellen 28 und 29 (Swiss Economics, 2018, S. 102f) ausgewiesenen Ergebnisse einen signifikanten Einfluss der untersuchten Parameter aufzeigen. Die Nicht-Berücksichtigung dieser Parameter erfolgt offenbar aufgrund eines zu geringen Einflusses (Impact). Dabei ist aber unklar, worauf dieses Kriterium beruht und weshalb 5% die kritische Einflussgrenze darstellen soll. Ebenso kann ein insgesamt geringer „Impact“ bei einzelnen Unternehmen durchaus große Auswirkungen haben.

Vor allem soll jedoch mit dem „optischen Eindruck“ der Abbildung 32 belegt werden, dass das Verhältnis von Zähl- zu Anschlusspunkten keine Rolle spielen würde oder ein „City Effekt“ ausgeschlossen werden könne (Swiss Economics, et al 2018, S. 103f). Die Analysemethode der „Draufsicht“ auf eine Abbildung ist jedoch kaum geeignet, einen „City Effekt“ auszuschließen.

Vergleicht man hingegen die Auswirkungen der Änderung des Normierungsparameters von Anschlusspunkten in RP 2 zu Zählpunkten kommt man zu anderen Schlussfolgerungen: Die Normierung mit Zählpunkten führt im aktuell vorliegenden Modell zu messbar schlechteren (bestabgerechneten) Effizienzwerten der Netzbetreiber mit einem hohen Verhältnis von Zähl- zu Anschlusspunkten, 

, als eine Normierung mit Anschlusspunkten.



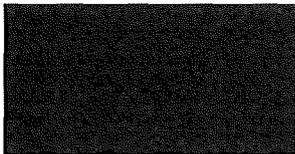
Forderung

Sensitivitätsanalysen, welche die Ergebnisse verschiedener Effizienzmodellrechnungen miteinander vergleichen, sind zur Modellvalidierung geeigneter als Second-Stage-Analysen.

Eine mögliche Benachteiligung von Unternehmen mit städtischem Versorgungsgebiet muss jedoch genauso ausgeschlossen werden wie die Benachteiligung von Netzbetreibern mit anderen heterogenen Versorgungsaufgaben.

Im Ergebnis möchten wir Sie um eine erneute Prüfung und Anpassung der Ausgestaltung des Effizienzwerts unter Berücksichtigung der obigen Ausführungen bitten.

Mit freundlichen Grüßen



Anlage