



Ermittlung des generellen sektoralen Produktivitätsfaktors für die vierte Regulierungsperiode Strom und Gas

Autoren:

Zweiter Entwurf

WIK-Consult

Dr. Marcus Stronzik
Dr. Bernd Sörries
Matthias Wissner
Dajan Baischew

Swiss Economics

Dr. Urs Trinkner
Dr. Lilia Ruslanova
Andreas Stritt

Impressum

WIK-Consult GmbH
Rhöndorfer Str. 68
53604 Bad Honnef
Deutschland
Tel.: +49 2224 9225-0
Fax: +49 2224 9225-63
E-Mail: info@wik-consult.com
www.wik-consult.com

Vertretungs- und zeichnungsberechtigte Personen

Geschäftsführerin	Dr. Cara Schwarz-Schilling
Direktor	Alex Kalevi Dieke
Direktor Abteilungsleiter Netze und Kosten	Dr. Thomas Plückebaum
Direktor Abteilungsleiter Regulierung und Wettbewerb	Dr. Bernd Sörries
Leiter der Verwaltung	Karl-Hubert Strüver
Vorsitzender des Aufsichtsrates	Dr. Thomas Solbach
Handelsregister	Amtsgericht Siegburg, HRB 7043
Steuer-Nr.	222/5751/0926
Umsatzsteueridentifikations-Nr.	DE 329 763 261

Inhaltsverzeichnis

Abbildungsverzeichnis	II
Tabellenverzeichnis	II
1 Einleitung	1
2 Methodische Ansätze zur Messung des technischen Fortschritts	2
2.1 Überprüfung des Stands der Wissenschaft	2
2.2 Darstellung und Bewertung anwendungsrelevanter methodischer Ansätze	11
2.3 Überprüfung der bisherigen methodischen Ausgestaltung des Produktivitätsfaktors	29
2.4 Zwischenfazit	36
3 Ermittlung der gesamtwirtschaftlichen Bestandteile nach § 9 ARegV	37
3.1 Stand der Wissenschaft	38
3.2 Annahme der Wettbewerbsfähigkeit	42
3.3 Zwischenfazit	45
4 Törnqvist-Methode	37
4.1 Daten	47
4.2 Methodik	48
4.3 Zwischenfazit	68
5 Malmquist-Methode	69
5.1 Grundansatz	69
5.2 Anwendung durch die Bundesnetzagentur in der RP3	70
5.3 Streitpunkte aus der RP3 und gerichtliche Beurteilung	71
5.4 Diskussion	75
6 Fazit	95
7 Literatur	96
8 Anhang: Berechnungen für den Gassektor anhand des Excel-Törnqvist-Tools	101
8.1 Datengrundlage	101
8.2 Technische Umsetzung	101
8.3 Weitere Plausibilisierungen	109

Abbildungsverzeichnis

Abbildung 2-1:	Methoden zur Ermittlung des Produktivitätsfortschritts	12
Abbildung 2-2:	Methode der kleinsten Quadrate	13
Abbildung 2-3:	Korrektur einer Produktionsschätzung bei COLS und MOLS	16
Abbildung 2-4:	Grundstruktur SFA-Schätzgleichung	18
Abbildung 2-5:	Stochastische Effizienzgrenze	19
Abbildung 2-6:	Schematische Darstellung der Effizienzgrenze bei der DEA	21
Abbildung 2-7:	Mittelwertbildung von zusammengesetzten Indexreihen	32
Abbildung 2-8:	Entwicklung des Xgen im Stromsektor	34
Abbildung 3-1:	Entwicklung des Global-Competitiveness-Indexes für Deutschland	44
Abbildung 3-2:	Entwicklung des Indexes World Competitiveness Yearbook für Deutschland	45
Abbildung 4-1:	Qualitätsmerkmale von Daten	47
Abbildung 4-2:	Netzentgeltdeflatoren Strom (Index 2015 = 100)	56
Abbildung 4-3:	Optionen zur Abbildung der Entwicklung der Eigenkapitalzinsen	61
Abbildung 4-4:	Entwicklung der Eigenkapitalzinsen für verschiedene Optionen (Gas)	62
Abbildung 5-1:	Geordnete Supereffizienzwerte Modell RP3 Fall 1 Totex	93

Tabellenverzeichnis

Tabelle 2-1:	Elemente der EU-KLEMS-2019-Datenbank, die verwendet werden, um die Produktivitätsmessung für RIIO-2 zu bilden	7
Tabelle 2-2:	Einfluss von Skalenerträgen auf Vergleichbarkeit von Unternehmen bei der DEA	22
Tabelle 2-3:	Methoden zur Produktivitätsmessung	27
Tabelle 4-1:	Vor- und Nachteile verschiedener Outputfaktoren	54
Tabelle 4-2:	Vor- und Nachteile verschiedener Deflatoren	57
Tabelle 4-3:	Wirtschaftszweige im Arbeitskostenindex Produzierendes Gewerbe und Dienstleistungsbereich	59
Tabelle 4-4:	Entwicklung der Inputpreise in der Position „Eigenkapitalzinsen“	61
Tabelle 4-5:	Vor- und Nachteile der Varianten	63
Tabelle 4-6:	Mittelwertbildung der Summanden des Xgen	66
Tabelle 5-1:	Bislang in Effizienzvergleichen verwendete Vergleichsparameter	80
Tabelle 5-2:	Auswirkungen von ausgewählten Ereignissen	81
Tabelle 5-3:	Vergleich der Resultate für PFG4 VNB mit alternativen SFA-Spezifikationen	87

Tabelle 5-4:	Vergleich der Resultate für den Frontier-Shift p.a. Gas mit Halbnormal- und Exponentialverteilungsannahme des Ineffizienzterms der SFA (Datengrundlage 28.02.2023)	89
Tabelle 8-1:	Xgen-Werte für verschiedene Zeiträume I (in %)	109
Tabelle 8-2:	Xgen-Werte für verschiedene Zeiträume II (in %)	109
Tabelle 8-3:	Xgen-Werte für zukünftige Jahre (in %)	110

1 Einleitung

Der generelle sektorale Produktivitätsfaktor (X_{gen}) ist ein wesentliches Element der Anreizregulierung.

Die ökonomische Notwendigkeit des generellen sektoralen Produktivitätsfaktors beruht auf seiner Aufgabe als Korrektiv des zur Inflationierung der Kostenbasis genutzten Verbraucherpreisgesamtindex sowie in der Abbildung des sektoralen technischen Fortschritts, wodurch eine sachgerechte Imitation wettbewerblicher Mechanismen gewährleistet wird.

Für die vierte Regulierungsperiode, die für Betreiber von Gasversorgungsnetzen am 01.01.2023 und für Betreiber von Stromversorgungsnetzen am 01.01.2024 beginnt, wird der Produktivitätsfaktor nun zum zweiten Mal von der Bundesnetzagentur nach wissenschaftlichen Methoden ermittelt und festgelegt, nachdem er zuvor per Verordnung vorgegeben wurde.

Der X_{gen} wird gemäß § 9 Abs. 1 der Verordnung über die Anreizregulierung der Energieversorgungsnetze (ARegV) aus der Abweichung des netzwirtschaftlichen Produktivitätsfortschritts vom gesamtwirtschaftlichen Produktivitätsfortschritt und der gesamtwirtschaftlichen von der netzwirtschaftlichen Einstandspreisentwicklung ermittelt. Die Ermittlung hat gemäß § 9 Abs. 3 ARegV nach Maßgabe von Methoden, die dem Stand der Wissenschaft entsprechen, zu erfolgen.

In diesem Gutachten werden Verfahren aufgezeigt, wie der generelle sektorale Produktivitätsfaktor berechnet werden kann. Zu diesem Zweck wird in Abschnitt 2 zunächst auf die Rechtsprechung, Literatur und internationale Erfahrung zum Thema „Stand der Wissenschaft“ eingegangen. Anschließend werden verschiedene theoretische Ansätze beschrieben und deren Vor- und Nachteile aufgezeigt. In Abschnitt 3 wird dargestellt, wie die gesamtwirtschaftlichen Bestandteile des generellen sektoralen Produktivitätsfaktors durch die Anwendung der sogenannten Residualbetrachtung ermittelt werden können. Abschnitt 4 betrachtet die Verwendung der Törnqvist-Methode in der dritten Regulierungsperiode und zeigt Vor- und Nachteile verschiedener Ansätze innerhalb der Berechnungen auf. Abschnitt 5 befasst sich mit der Malmquist-Methode, insbesondere der Auswahl der Vergleichsparameter und den ökonometrischen Schätzverfahren. In Abschnitt 6 werden weitere Fragestellungen und Sensitivitätsanalysen aufgezeigt.

2 Methodische Ansätze zur Messung des technischen Fortschritts

In diesem Abschnitt werden die in Frage kommenden methodischen Ansätze dargestellt und diskutiert. Zunächst wird allgemein auf den Stand der Wissenschaft eingegangen, um dann konkrete Methoden zu erörtern und die bisherigen methodischen Ansätze zu diskutieren.

2.1 Überprüfung des Stands der Wissenschaft

Nach § 9 Abs. 3 ARegV hat die Ermittlung des Xgen anhand von Methoden zu erfolgen, die dem Stand der Wissenschaft entsprechen. Die Verordnungsbegründung zu § 9 Abs. 3 ARegV besagt außerdem: *„Zur Berechnung des Faktors stehen der Regulierungsbehörde unterschiedliche, international anerkannte Methoden zur Verfügung. Hierzu zählen insbesondere Methoden, die auf Indexzahlen beruhen (z. B. der Törnqvistindex) sowie der Malmquistindex.“*¹

Der Begriff „Stand der Wissenschaft“ unterliegt dabei keiner allgemeingültigen Definition. Eine Diskussion um diesen Begriff fand in den 1990er-Jahren im Bereich des Atomrechts statt. Lohse (1993) schreibt dazu: *„Stand der Wissenschaft ist somit die Summe derjenigen wissenschaftlichen Erkenntnisse, die für die Beurteilung des wissenschaftlichen Fortschritts notwendig sind.“*² Und an anderer Stelle: *„Nach alledem kann kein Zweifel daran bestehen, dass insbesondere die Billigung durch die wissenschaftliche Gemeinschaft den Stand der Wissenschaft markiert.“*³

In eine ähnliche Richtung geht die Aussage des Bundesverwaltungsgerichts. In einem Beschluss aus dem Jahr 2004 zu einem Streit um Nachtflugbeschränkungen am Flughafen Berlin-Tegel stellt es fest: *„Ein neuer Stand der Wissenschaft ist jedoch nicht erreicht, solange bisher anerkannte wissenschaftliche Aussagen kritisch hinterfragt und kontrovers diskutiert werden, ohne dass sich in der Forschung bereits ein neuer Grundkonsens abzeichnet.“*⁴

Wie sich im Folgenden zeigen wird, sind die bisher angewendeten Methoden zur Bestimmung des generellen sektoralen Produktivitätsfortschritts in Deutschland (insbesondere der Törnqvist- und der Malmquistindex) in diesem Sinne auch weiterhin als Stand der Wissenschaft anzusehen.

Eine Annäherung kann über die bisherige Rechtsprechung im Energiebereich, einen Literaturüberblick über den wissenschaftlichen Diskussionsstand sowie die internationale Regulierungspraxis erfolgen. Auf diese Aspekte wird im Folgenden detailliert eingegangen.

¹ Bundesrat (2007), S.48/49.

² Lohse (1994), S. 102.

³ Lohse (1994), S. 101.

⁴ Bundesverwaltungsgericht (2004).

2.1.1 Rechtsprechung im Energiebereich

Das Oberlandesgericht Düsseldorf hat im Verfahren zum Xgen der dritten Regulierungsperiode Gas (2018 bis 2022) im Hinblick auf den Stand der Wissenschaft festgestellt:

„Anders als im Gefahrenabwehrrecht i.w.S. (vgl. dazu die Beschwerdebegründung v. 12.7.2018, S. 36 ff.) kommt es bei der Berechnung des generellen sektoralen Produktivitätsfaktors damit nicht auf die bestmögliche ökonomische Methode an. Es reicht vielmehr aus, dass eine Methode und ihre konkrete Umsetzung in Übereinstimmung mit der Verordnungsbegründung als wissenschaftlich vertretbar angesehen werden können. Bei der Anwendung von § 9 ARegV als Rechtsnorm sind ferner rechtsstaatliche Grundsätze wie die Rechtssicherheit, die Vorhersehbarkeit und vor allem die Anwendbarkeit, also mit Blick auf die Gerichte die Justiziabilität zu beachten.“⁵

Der Bundesgerichtshof hat zur Festlegung des Xgen Gas für die dritte Regulierungsperiode darüber hinaus zum Maßstab, an dem die Regulierungsentscheidung zu messen ist, wie folgt ausgeführt:

„Der Maßstab hierfür [für eine erneute Prüfung des gefundenen Ergebnisses, Anm. der Autoren] ergibt sich unbeschadet der Verpflichtung der Bundesnetzagentur, Methoden anzuwenden, die dem Stand der Wissenschaft entsprechen, nicht allein aus dem Stand der wissenschaftlichen Erkenntnis. Vielmehr ist zu berücksichtigen, dass die Regulierungsbehörde der Regulierungsentscheidung nicht ausweichen darf. Sie darf dies - jedenfalls grundsätzlich - auch dann nicht tun, wenn sich nach den Maßstäben guter wissenschaftlicher Praxis zu einer bestimmten tatsächlichen Frage verlässliche Erkenntnisse nicht oder nur mit einem zu dem zu erwartenden Erkenntniszuwachs außer Verhältnis stehenden Aufwand erreichen lassen. Vielmehr ist - im Regulierungsrecht nicht anders als bei der Beurteilung von Tatfragen auf anderen Rechtsgebieten - rechtlich zu bestimmen, in welchem Umfang niemals vollständig zu vermeidende tatsächliche Unsicherheiten bei der Prognosegrundlage hinzunehmen sind und wie sie sich auswirken.“⁶

Ob die Entscheidung der Regulierungsbehörde für eine Methode sachgerecht erfolgt ist, ist demnach nicht nur am Stand der Wissenschaft zu bemessen. Vielmehr sieht der Bundesgerichtshof auch Raum für die Berücksichtigung von Praktikabilitätsabwägungen, was insbesondere vor dem Hintergrund eines rollierenden Regulierungssystems sachgerecht erscheint.

Des Weiteren hat der Bundesgerichtshof festgestellt: *„Die Bundesnetzagentur hat den Stand der Wissenschaft bei der Wahl der Methoden zur Ermittlung des Produktivitätsfaktors zugrunde zu legen; zu seiner Fortentwicklung ist sie nicht verpflichtet.“⁷* Dies bedeutet insbesondere, dass gegenwärtig in der Wissenschaft gebräuchliche Methoden dem Anspruch „Stand der Wissenschaft“ genügen, eine inhaltliche Diskussion über mögliche Weiterentwicklungen ist der Regulierungsbehörde nicht abzuverlangen.

⁵ Oberlandesgericht Düsseldorf (2019), Rn. 66.

⁶ Bundesgerichtshof (2021), Rn 24.

⁷ Bundesgerichtshof (2021), Rn 44.

Schließlich stellt das Gericht auch in Bezug auf den Umgang mit den nach der Data Envelopment Analysis (DEA) und der Stochastic Frontier Analysis (SFA) ermittelten Ergebnissen fest, dass die Regulierungsbehörde einen Ermessensspielraum bei der Aggregation dieser Ergebnisse besitzt, wenn dies dem Stand der Wissenschaft entspricht, um eine möglichst treffsichere „Abschätzung der sektoralen Produktivitätsentwicklung auszufüllen.“⁸

Der Stand der Wissenschaft muss somit auch immer in den jeweiligen Kontext eingebettet werden. Die Anwendung muss robust, aber auch praktikabel sein.

2.1.2 Literatur

Die Geburtsstunde des Xgen wird üblicherweise in der Regulierung der britischen Telekommunikations-, Gas- und Wasserwirtschaft Ende der 1980er-Jahre und ab den 1990er-Jahren und deren Privatisierung gesehen. Dort wurde die Anreizregulierung in Form eines Regulierungsrahmens für die Erlösobergrenze eingeführt, genannt „RPI-X“ bzw. genereller „CPI-X“. Die Idee ist, die Erlösobergrenze über die Dauer einer Regulierungsperiode mit einem Preisindex zu inflationieren (z. B. Retail-Price-Index oder Consumer-Price-Index), von dem ein Faktor X abgezogen wird, der die im gewählten Preisindex für den regulierten Sektor nicht abgebildeten relevanten Faktoren abbilden soll.

Seither wurde der X-Faktor in vielen Ländern in der Regulierung eingesetzt. Auch die wissenschaftliche Debatte hat sich in dieser Zeit entwickelt. Um den Stand der Wissenschaft wiederzugeben, wird daher an dieser Stelle auf neuere Veröffentlichungen eingegangen, die sich mit dem Thema des generellen sektoralen Produktivitätsfaktors beschäftigen.

In ihrem Standardwerk beschreiben Coelli et al. (2005) als Methoden zur Produktivitätsmessung u. a. sowohl den Törnqvist-Index⁹, als auch den Malmquist-Index auf Basis der DEA als auch der Stochastic Frontier Analysis (SFA).¹⁰ Diese Methoden erscheinen also aus theoretischer Sicht grundsätzlich für die Messung von Produktivitätsfortschrittsraten geeignet.

Oxera (2017a) sieht die Törnqvist- und Malmquist-Methodik ebenfalls grundsätzlich als geeignet an, verweist aber auf vermeintliche Schwächen bei der Umsetzung durch die Bundesnetzagentur in der dritten Regulierungsperiode.¹¹

Ajayi et al. (2018) sehen in ihrem Bericht für die britische Regulierungsbehörde OFGEM (Office of Gas and Electricity Markets) sowohl die Törnqvist-Methode als auch den Malmquist-Index (auf der Basis von DEA und SFA) als etablierte Methoden zur Berechnung von Produktivitätsfortschrittsraten im Strom- und Gassektor an.¹² Dies zeigt auch die in

⁸ Bundesgerichtshof (2021), Rn 122.

⁹ Coelli et al. (2005), S. 90.

¹⁰ Coelli et al. (2005), S. 291 ff.

¹¹ Oxera (2017a), S. 1.

¹² Ajayi et al. (2018), S. 9.

diesem Artikel enthaltene Literaturübersicht.¹³ Von den dort aufgeführten 16 Produktivitätsuntersuchungen im Strom- und Gasbereich seit 1996 wenden acht eine DEA, sechs eine SFA, zwei eine ökonometrische Kostenfunktion und eine die Törnqvist-Methode an.

Meyer et al. (2020) befassen sich mit der Frage, ob der Frontier-Shift, also die Verschiebung der Effizienzgrenze über die Zeit, mit einem TOTEX-Malmquist richtig abgebildet wird. Grundsätzlich wird die Anwendung des Malmquist-Index innerhalb einer „CPI-X“-Regulierung von den Autoren als adäquat angesehen. Die Anwendung eines TOTEX-Malmquist sehen sie grundsätzlich als geeignet und für zwei Fälle als unverzerrt an:¹⁴

„Unter welchen Bedingungen berechnet der TOTEX Malmquist den korrekten Frontier-Shift? Wir betrachten zwei Fälle. Der erste Fall befasst sich mit technischer und allokativer Ineffizienz. Der TOTEX-Malmquistindex ist unverzerrt, wenn die Effizienzgrenze von Unternehmen gesetzt wird, die entweder in beiden Perioden technisch und allokativ effizient sind, oder wenn die Ineffizienzen dieser Grenzunternehmen über die Zeit konstant bleiben. Der zweite Fall betrachtet Unternehmen, die mit unterschiedlichen Inputpreisen konfrontiert sind. Der TOTEX-Malmquistindex ist unverzerrt, wenn die Inputpreise entweder konstant bleiben oder sich für alle Unternehmen im gleichen Verhältnis ändern.“¹⁵

Die Annahme technischer Effizienz ist angemessen, wenn die Anreizregulierung schon länger in Kraft ist. Die Annahme allokativer Effizienz kann verletzt sein, da bei langfristigen Vermögenswerten nicht kurzfristig auf Preisänderungen reagiert werden kann. Allerdings ist der TOTEX-Malmquist dann immer noch unverzerrt, wenn das Verhältnis der Inputpreise konstant bleibt. Im zweiten Fall können die Inputpreise unterschiedlich sein, wenn ein Regulierungsregime zwei strukturell und geographisch unterschiedliche Gebiete beinhaltet, in welchen zum Beispiel die Löhne signifikant unterschiedlich schnell steigen.¹⁶ Meyer et al. (2020) zeigen im Sektor des verarbeitenden Gewerbes und Dienstleistungen („Manufacturing Industry and Services“) eine unterschiedliche Lohnentwicklung zwischen 2008 und 2016 in verschiedenen Bundesländern auf. Diese ist jedoch nicht aussagekräftig, da es sich um einen anderen Sektor handelt. Die Frage der Berücksichtigung von Lohnunterschieden zwischen verschiedenen Bundesländern wurde im Nachgang zu den Effizienzvergleichen der zweiten Regulierungsperiode vor dem Oberlandesgericht Düsseldorf geführt. Das Oberlandesgericht Düsseldorf kam in seinem Beschluss VI-3 Kart 121/14 zu EVG2 zum Schluss, dass die Nichtberücksichtigung des regionalen Lohnniveaus sachgerecht sei, unter anderem weil die Netzbetreiber unternehmerische Entscheidungsmöglichkeiten haben, einem hohen Lohnniveau im Versorgungsgebiet zu entgehen bzw. das eigene Lohnniveau zu beeinflussen.¹⁷ Auch im Strombereich bieten sich den Netzbetreibern genügend Möglichkeiten, um bei Bedarf regional unterschiedlichen Lohnkosten bzw. – präziser gesprochen – Wertgrenzprodukten auszuweichen, sodass das Verhältnis der Inputpreise zwischen den Netzbetreibern konstant

¹³ Ajayi et al. (2018), S. 22 ff.

¹⁴ Meyer et al. (2020), S. 95, eigene Übersetzung.

¹⁵ Grundsätzlich kommt es zu Verzerrungen, wenn sich die allokativer Effizienz signifikant über die Zeit und unterschiedlich über die Netzbetreiber ändert. Die genannten Punkte sind Voraussetzungen, dass dies nicht so ist.

¹⁶ Meyer et al. (2020), S. 96.

¹⁷ Vgl. Oberlandesgericht Düsseldorf (2017), S. 37f.

bleibt.¹⁸ Diese Überlegungen gelten auch in einem dynamischen Kontext. Insofern kann die Folgerung von Meyer et al. (2020), dass der TOTEX-Malmquist im Allgemeinen zur Berechnung des Xgen verwendet werden kann, geteilt werden.¹⁹

Insgesamt kann festgestellt werden, dass sowohl die Törnqvist- als auch die Malmquist-Methode grundsätzlich als Stand der Wissenschaft zur Produktivitätsmessung im Strom- und Gassektor angesehen werden können. Dies spiegelt sich auch in der praktischen Umsetzung im internationalen Kontext wider, wie im nachfolgenden Abschnitt erläutert wird.

2.1.3 Internationale Regulierungspraxis

Im Folgenden wird dargelegt, mit welchen Methoden der generelle sektorale Produktivitätsfaktor im internationalen Kontext berechnet wird. Die Auswahl der Methoden der jeweiligen Regulierungsbehörden ist ein weiterer Hinweis auf den Stand der Wissenschaft.

2.1.3.1 Großbritannien

Die aktuelle Regulierungsperiode in Großbritannien läuft seit dem 1. April 2021 unter dem Titel „RIIO-2“. Der Frontier-Shift für Strom- und Gasnetzbetreiber wird als Differenz zwischen der sog. „Ongoing Efficiency“ und den realen Inputpreissteigerungen bezeichnet und beträgt für alle Netzbetreiber 1,15 % pro Jahr für die CAPEX und 1,25 % für die OPEX.²⁰ Die Entscheidung stützt sich auf eine Studie von CEPA²¹, in der die Produktivitäten auf Basis der EU-KLEMS Datenbank²² berechnet wurden.

Für die Feststellung der „Ongoing Efficiency“ schlägt CEPA (2020b) grundsätzlich vier Methoden vor:²³

- Growth-Accounting-Analyse, basierend auf einer Überprüfung der EU-KLEMS-Datenbank, mit einer Beschreibung der Effizienzbereiche, die sich aus verschiedenen Ansätzen ergeben
- vorausschauende Produktivitätsprognosen für die britische Wirtschaft
- historische Performance der Unternehmen und
- branchenspezifische Treiber möglicher Produktivitätssteigerungen in den Gas- und Stromnetzen, z. B. als Ergebnis der Innovationsförderung, die die Netzwerkunternehmen, während RIIO-1 erhalten haben.

Die *Growth-Accounting-Analyse* mit Daten der EU-KLEMS-Datenbank bezieht sich dabei auf verschiedene Vergleichssektoren, unterschiedliche Produktivitätskennzahlen und die

¹⁸ Vgl. Agrell et al. (2016), S. 12.

¹⁹ Meyer et al. (2020), S. 96.

²⁰ Ofgem (2020), S. 47 ff.

²¹ CEPA (2020a).

²² Vgl. <https://euklems.eu>, zuletzt abgerufen am 03.02.2022.

²³ CEPA (2020b), S. 5.

jeweils längsten verfügbaren Zeiträume. Tabelle 2-1 gibt eine Übersicht über die analysierten Aspekte.

Tabelle 2-1: Elemente der EU-KLEMS-2019-Datenbank, die verwendet werden, um die Produktivitätsmessung für RIIO-2 zu bilden

Zeitperiode	<ul style="list-style-type: none"> • 1997-2016, die den längsten Datensatz darstellen, der vollständige Wirtschaftszyklen in der EU-KLEMS-Datenbank 2019 abdeckt. • 2006-2016, der verwendet wird, um die Auswirkungen zu zeigen, wenn nur der letzte vollständige Konjunkturzyklus berücksichtigt wird, der in der EU-KLEMS-Datenbank verfügbar ist.
Produktivitätskennzahlen	<ul style="list-style-type: none"> • Totale Faktorproduktivität (TFP) und Partielle Faktorproduktivität (PFP: Arbeit und Vorleistungen). • Bruttowertschöpfung für TFP- und PFP-Maßnahmen. • Implizite Zahlen für den Produktionswert sowohl für TFP- als auch für PFP-Maßnahmen.
Vergleichsindustrien	<ul style="list-style-type: none"> • Baugewerbe. • Ungewichteter Durchschnitt ausgewählter Branchen (Herstellung von Chemikalien und chemischen Produkten; Herstellung von Computern, elektronischen und optischen Produkten, Herstellung von Elektrogeräten, Herstellung von Transportmitteln; Bauwesen; Groß- und Einzelhandel: Reparatur von Kraftfahrzeugen und Motorrädern; Transport und Lagerung; Finanz- und Versicherungstätigkeiten). • Ungewichteter Durchschnitt ausgewählter Branchen ohne verarbeitendes Gewerbe (Bauwesen; Groß- und Einzelhandel: Reparatur von Kraftfahrzeugen und Motorrädern; Transport und Lagerung; Finanz- und Versicherungstätigkeiten). • Ungewichteter Durchschnitt aller Branchen (ohne Immobilien, öffentliche Verwaltung, Bildung, Gesundheits- und Sozialdienste). • Gewichteter Durchschnitt aller Branchen (ohne Immobilien, öffentliche Verwaltung, Bildung, Gesundheits- und Sozialdienste).

Quelle: CEPA (2020b), S.12/13, eigene Übersetzung.

Die Spannweite des generellen sektoralen Produktivitätsfaktors wird also zum einen durch einen synthetischen Index gewisser Branchen²⁴ und zum anderen mit dem Durchschnitt aller Branchen berechnet. Für die Berechnungen selbst verwendet CEPA den Törnqvist-Index, da somit den Empfehlungen der OECD zur Berechnung von Produktivitätsänderungen gefolgt wird.²⁵

²⁴ Ein solcher synthetischer Vergleichsindex wurde auch für die dritte Regulierungsperiode in Deutschland diskutiert (vgl. Liebe et al., S. 73 ff.). Dieser Ansatz wurde aber nach Kritik aus der Branche nicht weiter verfolgt.

²⁵ CEPA (2020a), S. 17.

Vorausschauende Produktivitätsprognosen für die *Gesamtwirtschaft* sind ebenso Teil der Betrachtungen. Sie beruhen auf glaubwürdigen Quellen für zukunftsgerichtete Produktivitätsschätzungen wie die Bank of England (BoE) und das Office for Budget Responsibility (OBR).²⁶

Für die *historische Performance der Unternehmen* geht CEPA auf den Artikel von Ajayi et al. (2018) ein. Die Anwendung eines DEA-Malmquist wird grundsätzlich als Stand der Wissenschaft angesehen, einzelne Aspekte bei der Umsetzung werden aber kritisiert, beispielsweise die Verwendung des RPI zur Deflationierung.²⁷

Schließlich werden *spezifische Produktivitätstreiber* aufgrund der in Großbritannien grundsätzlich im Regulierungssystem verankerten Finanzierung von Innovationen untersucht. Als Ergebnis wird eine Produktivitätssteigerung von 0,2 % als Ausfluss der Finanzierung von Innovationsmaßnahmen aus der vorangegangenen Regulierungsperiode berechnet.²⁸

Festzuhalten ist, dass sich die Regulierungsbehörde OFGEM aber letztlich für den ersten Ansatz entschieden hat (Festlegung auf Basis der Törnqvist-basierten Berechnung mit EU-KLEMS-Daten). Die zusätzliche Steigerung um 0,2 % jährlich aus den Innovationsmaßnahmen wurde aber ebenso in die Überlegungen einbezogen.²⁹

2.1.3.2 Niederlande

In den Niederlanden wird der generelle sektorale Produktivitätsfortschritt für Gas- und Stromnetzbetreiber auf gleiche Weise berechnet. Die Berechnung erfolgt vergangenheitsorientiert. Die niederländische Regulierungsbehörde ACM führt hierzu aus:

„Je repräsentativer die verwendeten vergangenen Erkenntnisse und je robuster der Schätzer ist, desto genauer wird die Schätzung der zukünftigen Produktivitätsänderung sein, die regionale Netzbetreiber erfahren könnten. Repräsentativ bedeutet hier, dass erwartet wird, dass die verwendeten vergangenen Realisierungen mit zukünftigen Realisierungen vergleichbar sind, beispielsweise weil die Umstände im Bewertungszeitraum mit denen in dem Zeitraum vergleichbar sind, für den die Änderung der Produktivität geschätzt wird. Robust bedeutet, dass die endgültige Schätzung der Produktivitätsänderung auf möglichst vielen repräsentativen Erkenntnissen aus der Vergangenheit basiert. Dadurch werden die Auswirkungen von Zwischenfällen oder Messfehlern auf die Produktivitätsänderung minimiert.“³⁰

Die Berechnung selbst erfolgt über die jährlichen Veränderungsraten der effizienten Kosten pro Outputseinheit. Die einzelnen Outputs werden dabei gewichtet und zu einer Größe zusammengefasst. Die Input- und Outputdaten beruhen auf Erhebungen bei den Unternehmen.

²⁶ CEPA (2020b), S. 8.

²⁷ CEPA (2020b), S. 17 ff.

²⁸ CEPA (2020b), S. 25.

²⁹ Ofgem (2020), S. 213.

³⁰ ACM (2021a), Rn. 269, eigene Übersetzung.

2.1.3.3 Österreich

In Österreich findet die Festlegung des generellen sektoralen Produktivitätsfortschritts generell auf konsensualer Basis statt. Für die letzte Regulierungsperiode (2018 bis 2022) wurde für die Gasverteilternetzbetreiber eine Entscheidung getroffen, die zwischen den Ergebnissen eines Branchengutachtens³¹ und eines Gutachtens für die Regulierungsbehörde E-Control³² lag.

Die Berechnungen des Branchengutachtens basierten auf der Schätzung einer Translog-Kostenfunktion für den Zeitraum 2002 bis 2013.³³ Der Produktivitätsfortschritt wird dabei anhand eines linearen und quadratischen Zeittrends berechnet.³⁴ Die Gutachter der Regulierungsbehörde verwendeten denselben Ansatz und führten zusätzlich eine Ausreißeranalyse durch.³⁵

Für die aktuelle Regulierungsperiode (2023 bis 2027) wurde für die Gasverteilternetzbetreiber ähnlich vorgegangen. In der Grundspezifikation wurden die realen bereinigten OPEX des Unternehmens zum Zeitpunkt t durch die gewichtete Netzlänge und die Anzahl der Zähler erklärt. Um den Xgen zu schätzen, wurde ein linearer und quadrierter Zeittrend in die Schätzgleichung mit aufgenommen.³⁶

Die Festlegung der Regulierungsbehörde sah schließlich einen Wert von 0,4 % für den Xgen vor, ein Mittelwert aus einem Produktivitätsfortschritt von 0 % auf der einen Seite sowie einem Mittelwert aus der unteren Bandbreite des Behördengutachtens und des Wertes der vorangegangenen Regulierungsperiode auf der anderen Seite.³⁷

Für die aktuelle Regulierungsperiode im Stromsektor (2019 bis 2023) wurden für die Entscheidung verschiedene Gutachten berücksichtigt. Der Xgen wurde schließlich auf 0,815 % pro Jahr festgelegt.

Methodisch wurde ein ähnlicher Ansatz wie im Gasbereich verfolgt. Da die Anzahl der Daten für eine SFA nicht hinreichend war, wurde eine Kostenfunktion mit Zeittrend herangezogen.³⁸ Stronzik und Wissner (2018) berechneten den sektoralen Produktivitätsfortschritt im Gutachten für die E-Control zusätzlich mit einer Malmquist-DEA.³⁹ Die Datenauswahl erstreckte sich dabei auf den längsten verfügbaren Zeitraum.

Für die zweite Regulierungsperiode (Gas) und die dritte Regulierungsperiode (Strom) wurden die Berechnungen zum generellen sektoralen Produktivitätsfaktor auf Basis volkswirtschaftlicher Indexreihen und der Törnqvist-Methode durchgeführt.⁴⁰

³¹ Gugler und Liebensteiner (2017).

³² Schmitt et al. (2017).

³³ Gugler und Liebensteiner (2017), S. 4.

³⁴ Gugler und Liebensteiner (2017), S. 3.

³⁵ Schmitt et al. (2017), S. 18 ff.

³⁶ Baischew et al. (2022), S.33.

³⁷ E-Control (2022), S. 36.

³⁸ E-Control (2018), S. 15 ff.

³⁹ Stronzik und Wissner (2018), S. 28 ff.

⁴⁰ E-Control (2012), S. 15 ff., E-Control (2013), S. 17 ff.

2.1.3.4 Schweden

Schweden hat seit 2012 eine Anreizregulierung für Stromnetzbetreiber und seit 2015 für Gasnetzbetreiber, welche von der schwedischen Regulierungsbehörde Energimarknadsinspektionen (EI) festgelegt wird. Bisher gab es drei Regulierungsperioden für Strom (2012–2015, 2016–2019 und 2020–2023) und zwei Regulierungsperioden für Gas (2015–2018 und 2019–2022).

Beim Strom wird zwischen lokalen Netzbetreibern und größeren Netzbetreibern (regionale sowie „Kernnetzbetreiber“) unterschieden.

Für lokale Netzbetreiber wird ein individuelles Produktivitätsziel gesetzt. Dieses wird anhand einer Data Envelopment Analysis (DEA) unter Annahme konstanter Skalenerträge bestimmt und auf die kontrollierbaren OPEX angewendet. Die ins Modell eingehenden Outputvariablen sind die Anzahl der Ausspeisepunkte, die Netzlänge, Höchstlast als „maximum power output against overhead network“ sowie die Anzahl der Umspannwerke.⁴¹ In der aktuellen Regulierungsperiode (2020–2023) beträgt die Untergrenze für dieses Produktivitätsziel 1 % p.a. und die Obergrenze 1,82 % p.a.⁴²

Für größere Netzbetreiber wird ein genereller Produktivitätsfortschritt verwendet, da für diese keine ausreichende Anzahl an Netzbetreibern zur Berechnung eines individuellen Fortschritts vorhanden ist.⁴³ Dieser generelle Produktivitätsfortschritt für größere Netzbetreiber wurde für die Regulierungsperiode (2012–2015) aufgrund einer Analyse internationaler Erfahrungen und einer Produktivitätsanalyse des schwedischen Elektrizitätssektors auf 1 % p.a. festgelegt. Für Letztere hatte EI mit den Daten der Verteilnetzbetreiber aus den Jahren 2001–2008 ein jährliches Produktivitätswachstum von 2 % p.a. geschätzt.⁴⁴

Während die Berechnung des individuellen Produktivitätsfortschritts für die zweite und dritte Regulierungsperiode mit aktualisierten Basisjahren und Daten angepasst wurde, wird für den generellen Produktivitätsfortschritt weiterhin 1 % p.a. verwendet.⁴⁵

2.1.3.5 Norwegen

Das Norwegian Water Resources and Energy Directorate (NVE) reguliert den norwegischen Elektrizitätssektor seit 1997. Bis 2007 wurde eine RPI-X-Regulierung gewählt, danach wurde auf einen Yardstick-Regulierungsansatz gewechselt. Die Yardstick-Regulierung gibt die Erlösobergrenze vor, welche zu 40 % aus den tatsächlichen Kosten des Netzbetreibers C und zu 60 % aus den Kosten eines durchschnittlich effizienten Netzbetreibers C* bestehen.⁴⁶ Das Modell hat somit keinen separaten Term für den X-Faktor,

⁴¹ Energimarknadsinspektionen (2022), S. 38, eigene Übersetzung.

⁴² Energimarknadsinspektionen (2022), S. 50, eigene Übersetzung.

⁴³ Energimarknadsinspektionen (2022), S. 38, eigene Übersetzung.

⁴⁴ Schweinsberg et al. (2011), S. 51.

⁴⁵ Energimarknadsinspektionen (2022), S. 50, eigene Übersetzung.

⁴⁶ NVE (2021).

da dieser prozessendogen ist. Der Frontier-Shift zeigt sich in der Abnahme von C^* , den Kosten eines durchschnittlich effizienten Netzbetreibers.⁴⁷

Zur Berechnung von C^* wird ein dreistufiger Kostenvergleich verwendet. In der ersten Stufe werden mithilfe eines DEA-Modells der TOTEX firmenspezifische Effizienzwerte ermittelt. In der zweiten Stufe werden diese durch eine „Second-stage-regression“ mit Umweltvariablen bereinigt, so dass geographische und regionale Unterschiede miteinbezogen werden. In der dritten Stufe werden die Resultate so kalibriert, dass daraus die branchendurchschnittlichen Normkosten C^* gebildet werden können.⁴⁸ Die Berechnung der C^* wird genauso wie die Berechnung der C jedes Jahr anhand der Daten des vorvergangenen Jahres ($t-2$) durchgeführt.⁴⁹

Im Jahr 1997 wurde der generelle sektorale Produktivitätsfaktor auf 2 % p.a. festgelegt. Im folgenden Jahr wurde der Faktor auf 1,5 % p.a. gesenkt und auch für die restlichen Jahre der CPI-X-Regulierung so beibehalten. Da der generelle Produktivitätsfaktor in der Yardstick-Regulierung prozessendogen ist, muss dieser nicht explizit definiert werden, sondern wird durch die Abnahme der Normkosten C^* fortlaufend angepasst und in die Erlösbergrenze miteingerechnet.⁵⁰

2.2 Darstellung und Bewertung anwendungsrelevanter methodischer Ansätze

Die im vorherigen Abschnitt bereits erwähnten Methoden werden an dieser Stelle beschrieben und ihre Vor- und Nachteile aufgeführt. Anschließend werden sie im Hinblick auf verschiedene Kriterien gegenübergestellt.

Für die allgemeine rechtliche und regulatorische Einordnung des generellen sektoralen Produktivitätsfortschritts sei auf das Gutachten von Liebe et al. (2017, S. 18 ff.) verwiesen. An dieser Stelle konzentriert sich die Untersuchung auf die verschiedenen ökonomischen Methoden zur Produktivitätsmessung.

Grundsätzlich können verschiedene Methoden zur Messung des Produktivitätsfortschritts Anwendung finden. Ausgehend von Coelli et al. (2005, S.6) sind dies im Wesentlichen vier Ansätze, die entsprechend eingeordnet werden können (vgl. Abbildung 2-1): Ökonometrische Produktionsmodelle nach der Kleinste-Quadrate-Methode (OLS-Schätzungen), TFP-Indexzahlen, Data Envelopment Analysis (DEA) und Stochastic Frontier Analysis (SFA).

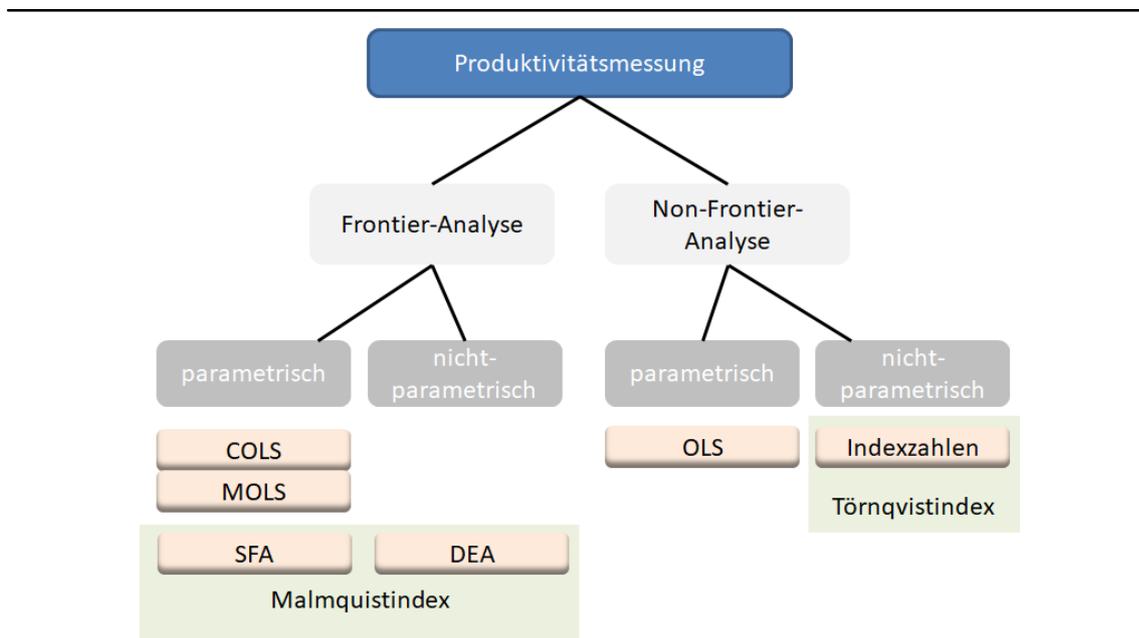
⁴⁷ Schmitt und Stronzik (2015), S. 33.

⁴⁸ Schmitt und Stronzik (2015), S. 34.

⁴⁹ Schweinsberg et al. (2011), S. 37.

⁵⁰ Schmitt und Stronzik (2015), S. 35f.

Abbildung 2-1: Methoden zur Ermittlung des Produktivitätsfortschritts



Quelle: Eigene Darstellung auf Basis von Coelli et al. (2005), S. 6, Parman und Featherstone (2019), S. 60ff.

Die grün hinterlegten Felder stellen dabei die Methodenauswahl der Bundesnetzagentur für die dritte Regulierungsperiode dar.

Die Methoden lassen sich unterscheiden in solche, die explizit eine Grenze des technischen Fortschritts berechnen (Frontier-Analyse) und solche, die dies nicht tun (Non-Frontier-Analyse).⁵¹ Dabei kann in parametrische und nichtparametrische Verfahren unterschieden werden. Parametrische Verfahren zeichnen sich dadurch aus, dass sie bis auf einige unbekannte Parameter, die aus den Daten geschätzt werden, vorher definiert werden. Nichtparametrische Modelle sind a-priori weniger restriktiv. Es werden nur sehr breit gefasste Eigenschaften angenommen.⁵²

Im Folgenden werden die Methoden einzeln dargestellt und auf jeweilige Vor- und Nachteile eingegangen.

2.2.1 Schätzung mit Zeittrend (OLS)

Eine Möglichkeit zur Bestimmung der Produktivitätsentwicklung besteht mittels einer Regression der Kostenfunktion nach der Methode der kleinsten Quadrate (Ordinary Least Squares, OLS) unter zusätzlicher Berücksichtigung eines Zeittrends.⁵³ Es handelt sich

⁵¹ Parman und Featherstone (2019), S. 61.

⁵² Bogetoft und Otto (2011), S. 17.

⁵³ Vgl. z. B. Polynomics und Jacobs University (2016), Gugler und Liebensteiner (2016, 2017) und Liebe et al. (2017).

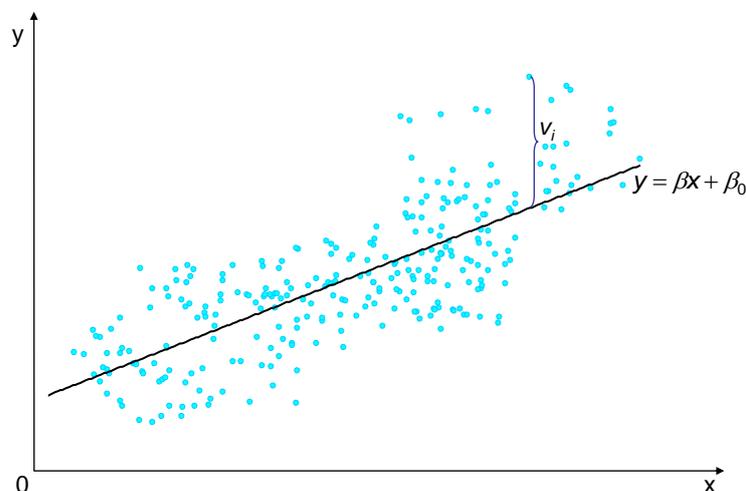
bei der OLS mit Zeittrend um eine Durchschnittsbetrachtung, es wird also keine Produktionsgrenze (Frontier) im eigentlichen Sinne ermittelt.⁵⁴

Die Methode dient dazu, den Wert von zu schätzenden Parametern (d. h. den Regressionsvariablen) so zu bestimmen, dass die Summe der quadrierten Abweichungen minimiert wird.⁵⁵ Eine weit verbreitete Anwendung stellt die lineare Regressionsanalyse dar.

Wie in Abbildung 2-2 zu sehen ist, gibt die (lineare) Schätzgleichung letztlich einen Erwartungswert für das durchschnittliche Verhalten aller Punkte/Unternehmen in den untersuchten Datengruppen wieder – die Schätzgerade verläuft durch die Mitte der Punktwolke, was aus dem Ansatz der Kleinstquadratschätzungen folgt, der positive und negative Abweichungen auszugleichen sucht. Die Abweichungen einzelner Unternehmen von den im Mittel erwarteten Werten werden in den Residuen v_i zusammengefasst, deren Summe einen Erwartungswert von Null hat, da sich im Mittel wiederum die positiven und negativen Abweichungen gerade aufheben.

Die lineare Schätzgleichung in Abbildung 2-2 beschreibt die Beziehung zwischen einem Input x und einem Output y , wobei sich die Analyse problemlos auf mehrere Inputs erweitern ließe.

Abbildung 2-2: Methode der kleinsten Quadrate



Aus Sicht einer Analyse, die damit befasst ist, die Effizienz einzelner Unternehmen oder eines ganzen Sektors zu untersuchen, besteht das Problem nun darin, dass die OLS-

⁵⁴ Die Ausführungen sind entnommen aus Franz und Stronzik (2005), S. 32 ff.

⁵⁵ Derart ermittelte Werte für die Regressionskoeffizienten sind sog. beste, lineare, erwartungstreue Schätzer, d. h. es lässt sich beweisen, dass es keinen linearen erwartungstreuen Schätzer gibt, dessen Varianz kleiner ist. In der englischsprachigen Literatur ist es üblich, solche Schätzer BLUE zu nennen (BLUE = best linear unbiased estimator).

Schätzgleichung von ihrer Konzeption her eben gerade nicht zum Ziel hat, die bestmögliche Performance abzubilden. Im Gegenteil vermischen sich aus Sicht einer an der Effizienz orientierten Regulierungsaufgabe Best und Worst Practice. Folge einer solchen Durchschnittsbetrachtung muss sein, dass etwa 50 % der Unternehmen mittelfristig gezwungen sind, sich dem Durchschnitt anzunähern. Dagegen bleibt die andere Hälfte der Unternehmen unbehelligt, obwohl die Analyse keinen Hinweis darauf geliefert hat, dass es sich bei dieser zweiten Hälfte ausschließlich um effiziente Unternehmen handelt.

Die hier verborgenen Potenziale können also nicht wohlfahrtssteigernd genutzt werden. Die dynamische Wirkung eines solchen Vorgehens ist zudem mit Zweifeln behaftet, da nicht ausgeschlossen werden kann, dass diejenigen Unternehmen, die besser als der Durchschnitt liegen, durch einen an einer OLS-Schätzung geleiteten Regulierungsansatz mit nachteiligen Anreizen konfrontiert sind. Die regulatorische Unsicherheit, nicht zu wissen, wie man sich relativ zum Sektor stellt, schwindet. Es ist anzunehmen, dass sich Unternehmen als Reaktion auf die regulatorische „Bestätigung“ ihrer relativen Güte in der Folge weniger effizienzorientiert zeigen. Dann aber dürfte die Punktwolke im Zeitablauf enger zusammenrücken, d. h. die v_i würden kleiner, jedoch würde sich die Regressionsgerade nicht bewegen. Im Gegensatz zur Gesamtwirtschaft würde die Produktivitätsentwicklung im Sektor also stagnieren.

Im Gegensatz zur SFA (vgl. Abschnitt 2.2.5) wird kein Term für die Ineffizienz berücksichtigt, so dass allein Annahmen über die funktionale Form der Schätzgleichung und die Verteilung des stochastischen Fehlerterms erforderlich sind.

Die dynamische Perspektive wird über die Integration eines Zeittrends in die Schätzgleichung erreicht. Über die Ableitung der Schätzfunktion nach der Zeit ($-\frac{\partial \ln C}{\partial t}$) kann die Produktivitätsentwicklung über die Zeit ermittelt werden. Dabei kann jedoch nicht mehr zwischen den einzelnen Ursachen einer Produktivitätsänderung differenziert werden. Es handelt sich vielmehr um die Bestimmung der Gesamtproduktivitätsänderung, mithin die Änderung der durchschnittlichen totalen Faktorproduktivität. Bei der Modellierung eines linearen Zeittrends gibt der geschätzte Koeffizient des Zeittrends direkt an, wie sich die Produktivität von der einen auf die andere Periode verändert hat.

Diesem Nachteil stehen jedoch auch einige Vorteile gegenüber. So ist der Ansatz aufgrund der Durchschnittsorientierung weniger anfällig gegenüber Ausreißern als z. B. die Malmquist-SFA. Die Nichtberücksichtigung des Ineffizienzterms umgeht, insbesondere wenn bei der SFA diesbezüglich eine exponentielle Verteilung gewählt wird (vgl. hierzu Abschnitt 5.4.6), auch mögliche Konvergenzprobleme, was zudem tendenziell eine deutlich höhere Robustheit der Ergebnisse impliziert. Darüber hinaus ist der OLS-Schätzer auch relativ robust gegenüber möglichen Annahmeverletzungen.

Das Datenerfordernis ist bei der OLS mit Zeittrend signifikant geringer als bei der Malmquist-SFA. Bei gleicher Kostenfunktion und gleicher Beobachtungszahl sind deutlich weniger Koeffizienten als bei der Malmquist-SFA zu bestimmen. Es verbleiben daher erheblich mehr Freiheitsgrade, was zu robusteren Schätzergebnissen führen kann. Bereits kleinere Stichproben können zu aussagekräftigen Ergebnissen führen.

2.2.2 Corrected Ordinary Least Squares (COLS)⁵⁶

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

- Zunächst wird eine einfache OLS-Schätzung durchgeführt, von deren Ergebnis bezüglich der Steigungsparameter β_j , mit $j = 1, \dots, k$, wobei k für die Anzahl der in die Schätzgleichung einbezogenen Erklärungsvariablen (Outputs) steht, angenommen wird, dass es auch im Hinblick auf die effiziente Grenze Gültigkeit hat. Das heißt, auch für das Grenzunternehmen kann nur angegeben werden, wie sich seine Produktion oder seine Kosten durchschnittlich entwickeln. Damit wird praktisch unterstellt, dass das effiziente Unternehmen mit einer ähnlichen Technologie arbeitet wie alle anderen Unternehmen des Sektors und lediglich bessere TFP-Werte erreicht.
- In einem zweiten Schritt wird dann der aus Sicht einer Frontier-Analyse verzerrte Schätzer des Achsenabschnitts der Regressionsfunktion β_0 korrigiert. Hierbei spielt die Frage der grundsätzlich unterstellten Verhaltensannahme (Kostenminimierung vs. Produktionsmaximierung) eine entscheidende Rolle, denn sie ist entscheidend für die Richtung der erforderlichen Korrektur. Bei Kostenminimierung wird die Regressionsgerade nach unten korrigiert, bei Produktionsmaximierung hingegen nach oben verschoben.

Bezogen auf eine Produktionsschätzung, wie sie Abbildung 2-3 zeigt, wird der Schätzer des Achsenabschnitts dann so ermittelt, dass

$$(2-1) \quad \beta_0^* = \beta_0^{OLS} + \max_i(v_i^{OLS})$$

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

$$(2-2) \quad -v_i^* = v_i^{OLS} - \max_i(v_i^{OLS})$$

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

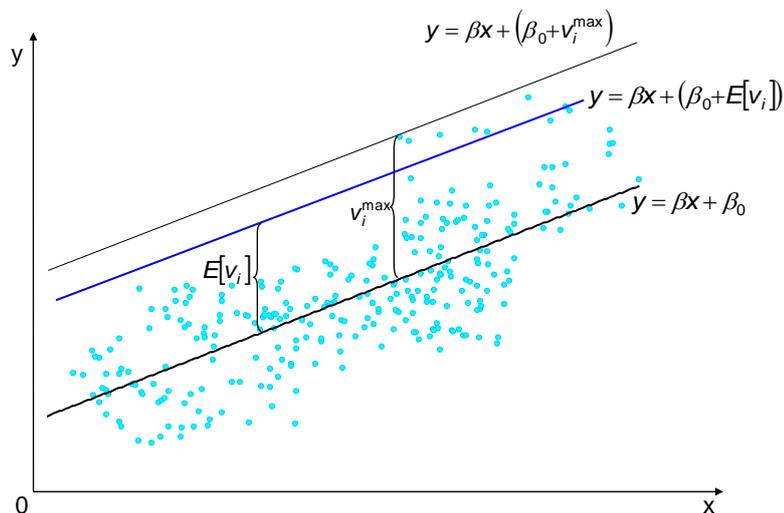
Geht man von normalverteilten v_i aus, d. h. von einem arithmetischen Mittel von Null, so resultiert aus diesem Vorgehen eine neue Normalverteilung, deren Mittelwert dann $-\max_i(v_i^{OLS})$ beträgt.

Damit werden die gesamten Störterme als Ineffizienz beschrieben. Dagegen findet keine Beachtung, dass sich auch aus von den Unternehmen nicht zu verantwortenden Umständen Produktionsverschiebungen ergeben können. Diese werden von der

⁵⁶ Die Ausführungen sind entnommen aus Franz und Stronzik (2005), S. 34 ff.

COLS in vollem Umfang als Ineffizienz interpretiert – es handelt sich insofern eigentlich nicht um ein stochastisches, sondern um ein deterministisches Modell.

Abbildung 2-3: Korrektur einer Produktionsschätzung bei COLS und MOLS



wik 

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

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

2.2.3 Modified Ordinary Least Squares (MOLS)⁵⁷

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

So verhält sich z. B. die Wahrscheinlichkeitsdichtefunktion einer exponentialverteilten Zufallsvariablen v entsprechend Gleichung (2-3): ⁵⁸

$$(2-3) f(v) = \frac{1}{\sigma_v} \cdot \exp\left\{-\frac{v}{\sigma_v}\right\}$$

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

$$(2-4) \beta_0^* = \beta_0^{OLS} + E(v_i) \text{ mit } E(v_i) \geq 0$$

$$(2-5) -v_i^* = v_i^{OLS} - E(v_i)$$

Auch die MOLS-Methode kann relativ einfach angewandt werden, sie beinhaltet jedoch das Risiko, dass nicht alle Unternehmen von der ermittelten Grenze eingeschlossen werden (angedeutet durch die mittlere Linie in Abbildung 2-3, bei der einige Beobachtungen über der Geraden liegen). Es kann nämlich nicht ausgeschlossen werden, dass für einzelne Beobachtungen $[v_i^{OLS} - E(v_i)] > 0$ gilt – es also Effizienzbewertungen gibt, die oberhalb von 1 liegen. Möglich ist auch, dass gar kein Unternehmen die effiziente Grenze erreicht, die wiederum durch die Parameter der OLS-Schätzung geprägt ist, weshalb alle gegen die COLS in diesem Zusammenhang geäußerten Kritikpunkte auch für die MOLS gelten müssen. Vor allem die Frage des Nichtunterscheidens zwischen zufälligen Ereignissen und Ineffizienzen bleibt problematisch. Hinzu kommt, dass die Verteilungsannahmen über die v_i unbegründet bleiben, weshalb unserer Ansicht nach auch die MOLS aus regulatorischer Sicht einen zu großen Unsicherheitsfaktor darstellen.

⁵⁷ Die Ausführungen sind entnommen aus Franz und Stronzik (2005), S. 36 ff.

⁵⁸ Praktisch werden bereits die MOLS-Ansätze zumeist mit der Maximum-Likelihood Methode, einer asymptotischen Annäherung an den wahren Wert, gelöst.

2.2.4 Malmquist-Index mit DEA und SFA

Der Malmquist-Index wurde in der dritten Regulierungsperiode in Deutschland auf Basis der Stochastic Frontier Analysis (SFA) und der Data Envelopment Analysis (DEA) angewendet. Für die grundlegende Funktionsweise sei auf Liebe et al. (2017) verwiesen.

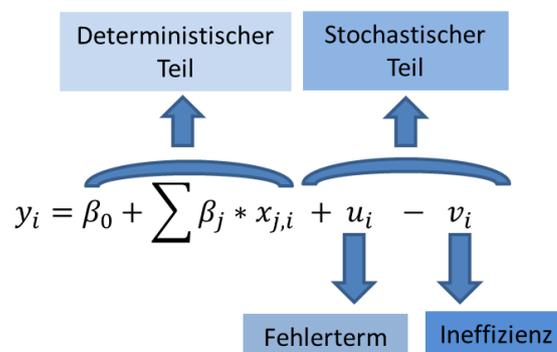
Der Malmquist-Index basiert auf der Grundidee, die Änderung von statischen Effizienzwerten (gemessen durch Input- oder Output-Distanz-Funktionen) in unterschiedlichen Perioden miteinander zu vergleichen und daraus die Produktivitätsentwicklung abzuleiten. Somit ist er ein Maß für die dynamische Effizienzentwicklung über die Zeit.

Wird ein Kosten-Malmquist angewendet, kann der ausgewiesene Frontier-Shift in den rein technologisch bedingten Frontier-Shift und einen Preiseffekt zerlegt werden.⁵⁹

2.2.5 Stochastic Frontier Analysis (SFA)

Während OLS, COLS und MOLS sog. deterministische Verfahren darstellen, fließt bei der SFA ein stochastischer Teil mit in die Schätzgleichung ein. Bei der SFA als parametrische Schätzung müssen also funktionale Zusammenhänge zwischen Inputs und Outputs bestimmt werden. Ebenso müssen Annahmen über zwei stochastische Elemente getroffen werden, welche in der Grundstruktur in Abbildung 2-4 erläutert werden.

Abbildung 2-4: Grundstruktur SFA-Schätzgleichung



wik 

Quelle: Liebe et al. (2017), S. 32.

In Abbildung 2-4 wird eine Produktionsfunktion mit dem Output y und den Inputs x angenommen.

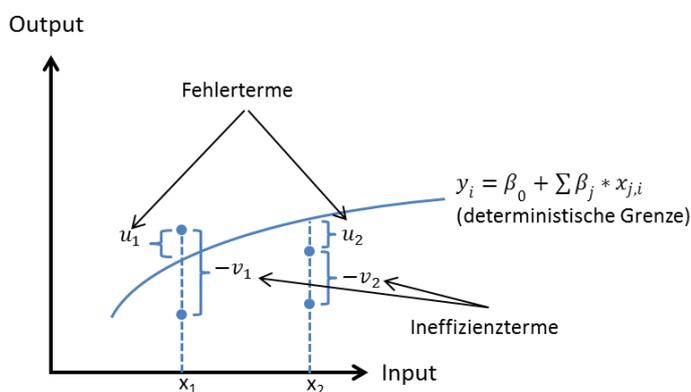
Die am weitesten verbreiteten Produktionsfunktionen in diesem Kontext ist die Cobb-Douglas-Funktion und die Translog-Funktion, welche eine generalisierte Form der Cobb-Douglas-Funktion darstellt und sich auf Grund ihrer Linearität in den Parametern besonders gut für ökonometrische Schätzungsmethoden eignet.

⁵⁹ Liebe et al. (2017), S. 38.

Die Einführung des normalverteilten Fehlerterms mit Erwartungswert Null führt zu einer symmetrischen Abweichung um die im deterministischen Teil der Gleichung festgelegten Effizienzgrenze. Ergebnis dessen ist eine stochastische Effizienzgrenze. Aufgrund von Ineffizienzen liegt die beobachtbare Outputmenge jedoch unterhalb dieser Grenze (schematische Darstellung in Abbildung 2-5).

Die technische Effizienz ist somit definiert als $e_{it} = u_{it} - v_{it}$, das heißt, Fehlerterm abzüglich der technischen Ineffizienz; der bedingte Erwartungswert von $\exp(-v_{it})$ kann herangezogen werden. Da v_{it} eine nicht negative, zufällige Variable ist, liegt die Vorhersage für die technische Effizienz zwischen Null und Eins, während ein Wert von Eins vollständiger technischer Effizienz entspräche.⁶⁰

Abbildung 2-5: Stochastische Effizienzgrenze



wik

Quelle: Liebe et al. (2017), S. 33..

Ein Vorteil der SFA ist die geringere Anfälligkeit gegenüber Ausreißern, verglichen mit der DEA, da ein stochastischer Teil mit in die Berechnung einfließt. Ebenso können Zeiteffekte besser von der SFA aufgenommen werden, da durch die Zeitvariablen explizit für zeitliche Effekte kontrolliert wird.

Die SFA mit ihrer hohen Flexibilität der Produktions- bzw. Kostenfunktionen erfordert jedoch eine Vielzahl an zu schätzenden Koeffizienten, was die Freiheitsgrade der Schätzung bedingt und somit eine hohe Anzahl an Beobachtungen fordert. Bei der SFA ist die Betrachtung des Ineffizienzterms und die Annahme über dessen Verteilung teilweise schwierig, da gerade bei kleinen Stichproben eine klare Trennung des Ineffizienzterms vom Fehlerterm nicht möglich ist. Dies geht mit dem Verlust der Separierbarkeit der unterschiedlichen Gründe für einen Produktivitätszuwachs einher.

⁶⁰ Coelli et al. (2005), S. 301.

Um der Problematik einer zu geringen Anzahl an Beobachtungen entgegen zu wirken, kann anstatt einer Panel-Analyse eine „pooled“-Variante angenommen werden. Bei dieser Variante wird die Zeitdimension vernachlässigt. Ebenso wird in einer „pooled“-Variante angenommen, dass alle Beobachtungen unabhängig sind.

2.2.6 Data Envelopment Analysis (DEA)

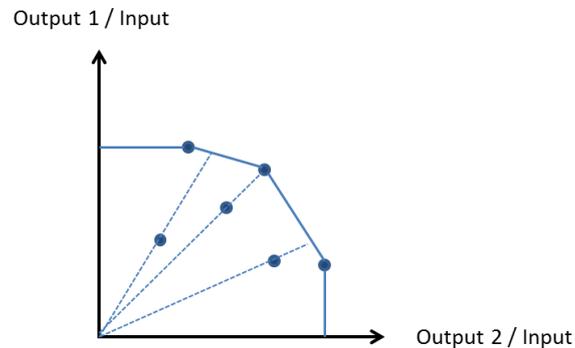
Bei der DEA werden Outputs und Inputs der jeweiligen Unternehmen ins Verhältnis gesetzt, sodass hinsichtlich der funktionalen Form zwischen Inputs und Outputs keine Annahmen getroffen werden müssen, das heißt, es muss keine explizite Produktionsfunktion angenommen werden. Die DEA beruht auf Methoden der linearen Optimierung bzw. Programmierung. Sie ist somit deterministisch und es fließen keine stochastischen Fehlerterme ein. Die Verhältniswerte die vom Input und Output gebildet werden, werden auf Werte zwischen Null und Eins bedingt, wobei ein Netzbetreiber mit einem Effizienzwert von Eins mit höchstmöglicher Effizienz arbeitet und somit die Effizienzgrenze abbildet. Da ein höchstmögliches Effizienzverhältnis aus unterschiedlichen Input- und Outputwerten entstehen kann, spricht man auch von Eckpunkten auf der Effizienzgrenze. Netzbetreiber mit einem Effizienzwert kleiner Eins erreichen ein weniger effizientes Verhältnis zwischen Input und Output (siehe Abbildung 2-6 mit jeweils einem Input und zwei Outputs).

Ein effizientes Verhältnis zwischen Input und Output wird erreicht, sobald technische und allokativer Effizienz sowie Effizienz durch Skaleneffekte vorliegen. Dabei misst die technische Effizienz die Fähigkeit eines Unternehmens, gegeben der Inputs, ein Maximum an Output zu erlangen. Die allokativer Effizienz beschreibt die Fähigkeit eines Unternehmens, die Inputs in einem optimalen Verhältnis gegeben ihrer Faktorpreise zu wählen. Durch einen optimalen Produktionsumfang bzw. Unternehmensgröße entsteht Effizienz durch Skaleneffekte. Dann produziert das Unternehmen mit konstanten Skalenerträgen. Produziert ein Unternehmen in einem nichtoptimalen Produktionsumfang, wird mit variablen Skalenerträgen produziert (steigenden oder fallenden).⁶¹

Unternehmen können ihren Effizienzwert optimieren, indem sie ihr Inputniveau bei gleichbleibendem Output senken, entweder durch den Einsatz effizienterer Technologien oder die Steigerung allokativer Effizienz durch Anpassungen des Inputeinsatzes an Preise der Faktormärkte. Ebenso kann eine Effizienzerhöhung durch die Erhöhung des Outputniveaus bei gegebenem Inputniveau erreicht werden, durch den Einsatz effizienterer Technologie, die Steigerung allokativer Effizienz durch Anpassung des Outputs an Preise der Absatzmärkte oder durch die Erhöhung des Produktionsumfanges.

⁶¹ Siehe Coelli et. al (2005), S. 52ff und S. 58.

Abbildung 2-6: Schematische Darstellung der Effizienzgrenze bei der DEA



Quelle: Liebe et al. (2017), S. 30.

Anhand der Unternehmen einer gegebenen Stichprobe können Effizienzgrenzen ermittelt werden. Dieses bedeutet jedoch, dass Unternehmen, die nicht in die Stichprobe fallen, aber Teil der Grundgesamtheit sind, gegebenenfalls oberhalb der ermittelten Effizienzgrenze liegen. Dies wird auch als Selektionseffekt bezeichnet, der bei allen stichprobenbasierten Verfahren auftreten kann. Ebenso können Netzbetreiber mit unplausiblen Input- und Outputwerten eine Grenze aufspannen, die für die übrigen Betreiber nicht realisierbar ist. Aufgrund von Mess- und Datenfehlern oder falschen Angaben können somit Ausreißer entstehen, gegenüber denen die DEA nicht robust ist. Um dem entgegenzuwirken, können Sensitivitätsanalysen vorgenommen werden, in der Ausreißer gesondert betrachtet oder völlig herausgenommen werden.

Ein Vorteil der DEA ist, dass keine explizite Produktions- oder Kostenfunktion angenommen werden muss, da lediglich Verhältnisse zwischen Input und Output untersucht werden. Einzig müssen Annahmen über Skalenerträge getroffen werden, das heißt Annahmen über das Verhalten der Outputmenge gegenüber einer Veränderung der Inputmengen. Diese werden in Tabelle 2-2 beleuchtet, wobei $f(x_1, x_2)$ eine beliebige Produktionsfunktion für einen Output mit zwei Inputs x_1 und x_2 darstellt und λ jeweils die Veränderung der Input- und Outputmengen.

Tabelle 2-2: Einfluss von Skalenerträgen auf Vergleichbarkeit von Unternehmen bei der DEA

	Annahme über Skalenerträge	Einfluss auf Effizienzwerte
$f(\lambda x_1, \lambda x_2) = \lambda f(x_1, x_2)$	Konstante Skalenerträge	Größe des Netzbetreibers hat keinen Einfluss
$f(\lambda x_1, \lambda x_2) \geq \lambda f(x_1, x_2)$	Nichtfallende Skalenerträge	Kleine Netzbetreiber werden nur mit kleinen Netzbetreibern verglichen, große hingegen mit allen
$f(\lambda x_1, \lambda x_2) \leq \lambda f(x_1, x_2)$	Nichtsteigende Skalenerträge	Große Netzbetreiber können nur mit großen Netzbetreibern verglichen werden, kleine hingegen mit allen

Quelle: Eigene Darstellung basierend auf Liebe et al., S. 31.

Sollte Heterogenität bei Skalenerträgen vorliegen, kann durch die Inklusion einer zusätzlichen Nebenbedingung dafür gesorgt werden, dass nur Unternehmen ähnlicher Größe miteinander verglichen werden.

2.2.7 Indexzahlen⁶²

Die Berechnung der Produktivitätsveränderung kann auf der Basis sogenannter Indexzahlen (Index-Numbers) erfolgen. Allgemein ergibt sich die Produktivität aus dem Quotienten eines Output- und eines Inputindex:

$$\text{Produktivität} = \frac{\text{Outputindex}}{\text{Inputindex}}$$

Wird eine partielle Faktorproduktivität gemessen, so wird der Inputindex nur aus einem Input (z. B. Arbeit) gebildet. Zur Messung der TFP werden sämtliche Inputs herangezogen. Sie gehen nach ihrer relativen Bedeutung in den Inputindex ein. Im Regulierungskontext interessiert oftmals nicht nur die Produktivität zu einem bestimmten Zeitpunkt, sondern die Veränderung dieser Größe über die Zeit. Soll eine Veränderung der Produktivität gemessen werden, so ergibt sich für die TFP:⁶³

$$(2-6) \quad \Delta TFP = \frac{TFP_t}{TFP_{t-1}} = \frac{\frac{\text{Outputindex}_t}{\text{Inputindex}_t}}{\frac{\text{Outputindex}_{t-1}}{\text{Inputindex}_{t-1}}} = \frac{\text{Outputindex}_t}{\text{Outputindex}_{t-1}} \cdot \frac{\text{Inputindex}_{t-1}}{\text{Inputindex}_t}$$

Die Veränderung der TFP lässt sich demnach aus dem Quotienten der Veränderung des Outputindex und der Veränderung des Inputindex zwischen zwei Zeitpunkten berechnen.

⁶² Dieser Abschnitt ist weitestgehend entnommen aus Niederprüm et al. (2020), S. 9 ff.

⁶³ Müller (2009).

Bei der Verwendung von Indexzahlen für die Produktivitätsmessung sind zwei wesentliche Punkte zu beachten:

- Grundsätzlich fließen in die Produktivitätsmessung Mengen ein (z. B. die Anzahl der verteilten kWh (Output) oder die Anzahl der Arbeitsstunden (Input)). In den meisten Fällen lassen sich diese aber nicht numerisch beziffern bzw. sind nicht verfügbar. Auch können sie, bei mehr als einem Input oder Output, nur schwer oder gar nicht mathematisch verknüpft werden. Aus diesem Grund wird auf Wertgrößen abgestellt, also z. B. die (wertmäßige) Veränderung des Kapitalstocks. Da jedoch nur die Mengen- und nicht die Preisveränderungen in die Produktivitätsmessung eingehen sollen, müssen die nominalen Wertgrößen durch die Verwendung geeigneter Deflatoren in reale Größen umgerechnet werden.
- Ein weiterer Punkt ist die Gewichtung von verschiedenen Inputs innerhalb der Berechnung des Inputindex bzw. von verschiedenen Outputs innerhalb der Berechnung des Outputindex, falls mehrere Größen in die jeweilige Berechnung eingehen. Die Gewichtung ergibt sich als prozentualer Anteil eines Parameters an der Gesamtgröße, z. B. inputseitig der Anteil der Arbeitskosten an den Gesamtkosten.

Ein Vorteil der Indexzahlen ist zunächst, dass sie einfach zu berechnen und leicht reproduzierbar sind. Somit können mögliche Fehlerquellen besser identifiziert und eliminiert werden. Eine Berechnung ist sowohl mit Unternehmensdaten als auch mit aggregierten Daten, etwa der volkswirtschaftlichen Gesamtrechnung (VGR), möglich. Letztere sind normalerweise gut verfügbar und robust gegenüber Datenunsicherheiten. Auch besteht durch die oft gute Verfügbarkeit die Möglichkeit, möglichst lange Zeitreihen zu nutzen, um mögliche Effekte durch Konjunktur- und Investitionszyklen zu glätten.

Ein Nachteil von Indexzahlen liegt in der fehlenden Möglichkeit, die Ursachen für Effizienzgewinne gesondert zu identifizieren (allokative, technische und Skaleneffizienz sowie technologischer Fortschritt). Ziel von Produktivitätsmessungen auf der Grundlage von Indexzahlen ist es mithin „nur“, den technologischen Fortschritt zu erfassen. Indexzahlen bilden den technologischen Fortschritt nur dann ab, wenn angenommen werden kann, dass allokative, technische und Skaleneffizienz gegeben sind und diese sich nicht über die Zeit verändern. Bei einer Verbesserung (Verschlechterung) dieser drei Effizienzkatégorien wird der technologische Fortschritt somit tendenziell überschätzt (unterschätzt). Auch besitzt die Verwendung langer Zeitreihen z. B. aus der VGR dann einen Nachteil, wenn Brüche in den Erhebungsdaten vorliegen, z. B. aufgrund einer neuen Erhebungssystematik.

Die Berechnung der Veränderung der TFP kann durch verschiedene Indizes erfolgen.⁶⁴ Im Folgenden werden der Laspeyres-, Paasche, Fisher- und Törnqvist-Index beschrieben und auf ihre Eignung für die Produktivitätsmessung analysiert. Sie unterscheiden sich im Wesentlichen durch die Wahl der Periode, auf die die Veränderung bezogen wird.

64 McLellan (2004).

Ein Vorteil aller dargestellten Verfahren ist, dass bereits auf Basis von nur zwei Datenpunkten eine Produktivitätsänderung gemessen werden kann.

2.2.8 Laspeyres-Index, Paasche-Index und Fisher-Index

Der Laspeyres- und der Paasche-Index sind weit verbreitet. Sie können als Mengen- oder Preisindex genutzt werden. Der Laspeyres-Preisindex wird beispielsweise durch das statistische Bundesamt zur Berechnung des allgemeinen Verbraucherpreisindex (VPI) verwendet.⁶⁵

Zur Berechnung der Produktivitätsveränderung werden beim **Laspeyres-Mengenindex** die Preise der Basisperiode (0) der verschiedenen Inputs (Inputindex) bzw. Outputs (Outputindex) als Referenz herangezogen. Es ergibt sich somit eine reine Veränderung der Mengen, während die Preise konstant auf dem Niveau der Basisperiode gehalten werden. Wenn also z. B. für den Inputindex Preise und Mengen bekannt sind, ergibt sich folgende Definition für den Laspeyres-Mengenindex:

$$(2-7) \quad I_L = \frac{\sum_{m=1}^M p_0^m q_t^m}{\sum_{m=1}^M p_0^m q_0^m} = \frac{\sum_{m=1}^M \frac{q_t^m}{q_0^m} q_0^m p_0^m}{\sum_{m=1}^M p_0^m q_0^m} = \sum_{m=1}^M \frac{q_t^m}{q_0^m} \cdot \frac{q_0^m p_0^m}{\sum_{m=1}^M p_0^m q_0^m} = \sum_{m=1}^M \frac{q_t^m}{q_0^m} \cdot w_0^m$$

Mit

p_0^m = Preis in der Basisperiode

p_t^m = Preis in der Berichtsperiode

q_0^m = Menge in der Basisperiode

q_t^m = Menge in der Berichtsperiode

Wie ersichtlich wird, stellt w_0^m den Anteil des Inputs m am gesamten Input in der Basisperiode dar. Dieser Anteil wird multipliziert mit der Mengenänderung des jeweiligen Inputfaktors. Die Summe aller auf diese Weise gebildeten Produkte ergibt den Laspeyres-Index.

Beim **Paasche-Mengenindex** werden zur Berechnung der Produktivitätsveränderung die Preise der Berichtsperiode (t) der verschiedenen Inputs (Inputindex) bzw. Outputs (Outputindex) als Referenz herangezogen. Es ergibt sich somit eine reine Veränderung der Mengen, während die Preise konstant auf dem Niveau der Berichtsperiode gehalten werden. Wenn also z. B. für den Inputindex Preise und Mengen bekannt sind, wird der Paasche-Mengenindex wie folgt berechnet:

$$(2-8) \quad I_P = \frac{\sum_{m=1}^M p_t^m q_t^m}{\sum_{m=1}^M p_t^m q_0^m} = \frac{\sum_{m=1}^M \frac{q_t^m}{q_0^m} q_0^m p_t^m}{\sum_{m=1}^M p_t^m q_0^m} = \sum_{m=1}^M \frac{q_t^m}{q_0^m} \cdot \frac{q_0^m p_t^m}{\sum_{m=1}^M p_t^m q_0^m} = \sum_{m=1}^M \frac{q_t^m}{q_0^m} \cdot w_t^m$$

Der Term w_t^m stellt den Anteil des Inputs m am gesamten Input in der Berichtsperiode dar. Dieser Anteil wird multipliziert mit der Mengenänderung des jeweiligen Inputfaktors. Die Summe aller auf diese Weise gebildeten Produkte ergibt den Paasche-Mengenindex.

⁶⁵ Statistisches Bundesamt (2018).

Der Mengenindex nach Laspeyres (Paasche) zeigt somit an, wie sich die Menge geändert hätte, wenn das in der Basisperiode (Berichtsperiode) geltende Preisniveau unverändert auch in der Berichtsperiode (Basisperiode) herangezogen wird.

Beide Indizes besitzen den Nachteil, dass sie ein festes Bezugsjahr für die Gewichtung benötigen. Somit schlagen sich Veränderungen der Inputpreise im Zeitablauf (beim Laspeyres-Index mit zunehmendem Abstand von der Basisperiode) nicht in einer (eigentlich notwendigen) Änderung der Gewichtung nieder. Dadurch kann es zu Verzerrungen kommen.

Der Paasche-Index zieht immer das Berichtsjahr als Gewichtungsgrundlage heran. Dies wiederum erschwert eine Vergleichbarkeit des Indexes zwischen den Perioden, da unterschiedliche Berichtsjahre und somit unterschiedliche Preisniveaus in die Berechnung eingehen. Der Laspeyres-Index unterschätzt den Produktivitätsfortschritt aufgrund des Heranziehens von Vergangenheitswerten tendenziell, während der Paasche-Index ihn tendenziell überschätzt.⁶⁶

Die Nachteile der beiden Indizes werden durch den **Fisher-Index** abgemildert, der im Folgenden diskutiert wird. Der Fisher-Index versucht die Nachteile des Laspeyres- und Paasche-Index aufzufangen, indem er das geometrische Mittel der beiden Indizes bildet. Für einen Inputindex nach Fisher ergibt sich somit:

$$(2-9) \quad I_F = \sqrt{I_L \cdot I_P} = \sqrt{\frac{\sum_{m=1}^M p_0^m q_t^m}{\sum_{m=1}^M p_0^m q_0^m} \cdot \frac{\sum_{m=1}^M p_t^m q_t^m}{\sum_{m=1}^M p_t^m q_0^m}}$$

Der Fisher-Index ist auch als „Fisher-Ideal-Index“ bekannt, da er eine Reihe an Eigenschaften erfüllt.⁶⁷ Diewert (1992) empfiehlt daher den Fisher-Index zur Produktivitätsmessung.⁶⁸

2.2.9 Törnqvist-Index

Der Törnqvist-Index gewichtet, wenn ein Inputindex berechnet wird, die verschiedenen Inputs jeweils zur Hälfte mit dem Anteil der Berichtsperiode und dem Anteil der Vorperiode. Über die so gewichteten Mengenänderungen der Inputs wird das Produkt gebildet, der Törnqvist-Mengenindex ergibt sich also wie folgt:

$$(2-10) I_T = \prod_{m=1}^M \left(\frac{q_t^m}{q_{t-1}^m} \right)^{\frac{1}{2} (w_{t-1}^m + w_t^m)}$$

⁶⁶ Müller (2009).

⁶⁷ Dazu gehören z. B. die Zeitumkehrbarkeit und die Faktorumkehrbarkeit. Nähere Erläuterungen dazu finden sich beispielsweise in: Coelli, Rao, Battese (2002).

⁶⁸ Diewert (1992).

Die Anteile in den Perioden w_{t-1}^m und w_t^m sind die Wertanteile des jeweiligen Inputs am Gesamtinput. Der Anteil berechnet sich für den Input m in der Periode t wie folgt:

$$(2-11) \quad w_t^m = \frac{p_t^m q_t^m}{\sum_{m=1}^M p_t^m q_t^m}$$

Der Törnqvist-Index wird in der Produktivitätsmessung sehr häufig angewendet, weil er einfach anwendbar ist und ihm mit der Translog-Funktion eine flexible Produktionsfunktion zu Grunde liegt.⁶⁵ Der Törnqvist-Index kann deswegen als sog. superlativer Index bezeichnet werden. Conrad (1985) folgert daraus, dass der Törnqvist-Index „den traditionellen Indizes wie Paasche, Laspeyres oder auch einem geometrischen Mittel oder letztendlich Fishers idealem Index überlegen ist.“

2.2.10 Gegenüberstellung der Methoden

Die in den vergangenen Abschnitten beschriebenen Methoden besitzen verschiedene Vor- und Nachteile.

Tabelle 2-3: Methoden zur Produktivitätsmessung

Methode	Annahmen	Erforderlicher Stichprobenumfang	Betrachtungszeitraum	Datenanforderungen und -verfügbarkeit	Gesamtbewertung
OLS	<ul style="list-style-type: none"> Sämtliche Abweichungen zur Regressionsgeraden werden als Ineffizienzen unterstellt Keine Unterscheidung in Catch-up und Frontier-Shift Durchschnittsbetrachtung 	<ul style="list-style-type: none"> Relativ gering im Vergleich zu beispielsweise SFA 	<ul style="list-style-type: none"> Möglichst lange Möglichst Zeitreihen über das gesamte Stützintervall für alle Input- und Outputvariablen 	<ul style="list-style-type: none"> Datenerfordernis relativ gering für robuste Schätzungen Unternehmensdaten vorteilhaft 	<ul style="list-style-type: none"> Deterministische Methode Keine Modellierung des Frontier-Shifts Keine greifbare Überlegenheit gegenüber anderen Methoden
COLS	<ul style="list-style-type: none"> Parallelverschiebung und Ausrichtung der ermittelten Regressionsgerade an entsprechenden Extremwerten. Abstand zwischen Messpunkten und Effizienzgerade wird als Ineffizienz definiert, ohne weitere Unterteilung⁶⁹ Effiziente Technologie hat dieselbe Struktur wie durchschnittliche Technologie 	<ul style="list-style-type: none"> Siehe OLS 	<ul style="list-style-type: none"> Siehe OLS 	<ul style="list-style-type: none"> Siehe OLS 	<ul style="list-style-type: none"> Deterministische Methode Keine Modellierung des Frontier-Shifts Starke Anfälligkeit gegenüber Ausreißern Keine greifbare Überlegenheit gegenüber anderen Methoden
MOLS	<ul style="list-style-type: none"> Es wird eine Fehlerverteilung bei der Effizienzmessung angenommen und die Effizienzlinie zugunsten der weniger produktiven Einheiten stochastisch modifiziert. 	<ul style="list-style-type: none"> Siehe OLS 	<ul style="list-style-type: none"> Siehe OLS 	<ul style="list-style-type: none"> Siehe OLS 	<ul style="list-style-type: none"> Deterministische Methode Keine Modellierung des Frontier-Shifts Keine greifbare Überlegenheit gegenüber anderen Methoden

⁶⁹ Vgl. Burger (2008), S. 49.

Methoden	Annahmen	Erforderlicher Stichprobenumfang	Betrachtungszeitraum	Datenanforderungen und -verfügbarkeit	Gesamtbewertung
	<ul style="list-style-type: none"> Sämtliche Abweichungen zur verschobenen Regressionsgeraden werden als Ineffizienzen interpretiert 				
DEA (Malmquist)	<ul style="list-style-type: none"> Annahme über Skalenerträge notwendig Bei Kostenmalmquist keine separate Berechnung der sektoralen Inputpreisänderungen notwendig 	<ul style="list-style-type: none"> Gering 	Mindestens zwei Perioden	<ul style="list-style-type: none"> Datenerhebungen im Rahmen des Effizienzvergleichs Geprüfte Netzbetreiberdaten 	<ul style="list-style-type: none"> Modellierung des Frontier-Shift
SFA (Malmquist)	<ul style="list-style-type: none"> A-priori-Annahme über Zusammenhänge und Verteilungsfunktion (Fehler / Rauschen) Bei Kostenmalmquist keine separate Berechnung der sektoralen Inputpreisänderungen notwendig 	<ul style="list-style-type: none"> Große Anzahl an Untersuchungseinheiten und Daten notwendig 	<ul style="list-style-type: none"> Mindestens zwei Perioden 	<ul style="list-style-type: none"> Datenerhebungen im Rahmen des Effizienzvergleichs Geprüfte Netzbetreiberdaten 	<ul style="list-style-type: none"> Modellierung des Frontier-Shift
Törnqvist-Index	<ul style="list-style-type: none"> Keine Unterscheidung in Catch-up und Frontier-Shift Verzerrungen bei Änderung der allokativen, technischen oder Skaleneffizienz Berechnung üblicherweise auf Basis Bruttowertschöpfung oder Produktionswert als Output Translog-Produktionsfunktion 	<ul style="list-style-type: none"> Möglichst Netzbetreiberindizes oder umfassende Unternehmenszeitreihen Mögliche Brüche in den Zeitreihen müssen beachtet werden 	<ul style="list-style-type: none"> Möglichst langes Stützintervall (Glättung längerer Investitionszyklen bei den Energienetzen aufgrund langlebiger Anlagengüter) 	<ul style="list-style-type: none"> Veröffentlichungen des Statistischen Bundesamtes und/oder Daten der Netzbetreiber Berechnung auf Basis Bruttowertschöpfung oder Produktionswert 	<ul style="list-style-type: none"> Frontier-Shift unter Annahme unveränderter technischer, allokativer und Skaleneffizienz Methode vergleichsweise einfach und nachvollziehbar International gebräuchlich

2.2.11 Fazit zur Methodenauswahl

Die Auswahl der geeigneten Methodik sollte vor dem Hintergrund der jeweiligen Gesamtkonstellation in Hinsicht auf die verfügbaren Daten, die Datenqualität und den Betrachtungszeitraum erfolgen. SFA und DEA sind explizite Methoden zur Bestimmung des Frontier-Shifts, während der Törnqvist-Index durch seine relativ einfache Berechnung und Nachvollziehbarkeit Vorteile besitzt.

2.3 Überprüfung der bisherigen methodischen Ausgestaltung des Produktivitätsfaktors

In diesem Abschnitt werden wesentliche Aspekte der bisherigen Vorgehensweise der Bundesnetzagentur aufgegriffen und diskutiert. Dabei werden wesentliche Diskussionspunkte, die in der Konsultationsphase und/oder in verschiedenen Gerichtsverfahren eine Rolle gespielt haben, beleuchtet, um entsprechende Transparenz für die mögliche Vorgehensweise für die vierte Regulierungsperiode herzustellen.

2.3.1 Regulatorische vs. handelsrechtliche Daten

Der Törnqvist-Index wurde für die dritte Regulierungsperiode auf einer handelsrechtlichen Basis berechnet, während der Malmquist-Index auf den kalkulatorischen Kosten der Effizienzvergleiche beruhte.

Der Hintergrund der Berechnungen auf handelsrechtlicher Basis war u. a. Kritik aus der Branche, nach der Berechnungen der Veränderung der Totalen Faktorproduktivität (TFP) auf Basis von Daten des statistischen Bundesamtes nicht getrennt für Strom und Gas möglich seien.⁷⁰ Die Bundesnetzagentur erhob daher entsprechende handelsrechtliche Daten bei den Netzbetreibern, getrennt nach Strom- und Gasnetzbetreibern.⁷¹ In der Methodik der Berechnungen wurde sich dann an die Vorgehensweise von Berechnungen mit Daten des statistischen Bundesamtes angelehnt. Die Daten konnten dabei im Wesentlichen den jeweiligen Jahresabschlüssen entnommen werden. Die Datenerhebung war somit für die Netzbetreiber weniger aufwendig als die umfangreichere fünfjährige Erhebung zu den Effizienzvergleichen, da diese auf letztlich im Bundesanzeiger veröffentlichten Daten oder deren Teilmengen beruhte. Auch wäre eine jährliche Kostenprüfung mit erheblich höherem Aufwand verbunden gewesen.

Die Bundesnetzagentur hat die Berechnungen an dieser Stelle sodann konsequent auf handelsrechtlichen Daten basiert, dies gilt z. B. auch für die Abschreibungen. Die einzige

⁷⁰ Oxera (2017b), S. 24.

⁷¹ Vgl. https://www.bundesnetzagentur.de/DE/Beschlusskammern/1_GZ/BK4-GZ/2017/BK4-17-0004/BK4-17-0004_Beschluss_Download_bf.pdf?_blob=publicationFile&v=2, zuletzt abgerufen am 10.02.2022 und https://www.bundesnetzagentur.de/DE/Beschlusskammern/1_GZ/BK4-GZ/2017/BK4-17-0094/BK4-17-0094_Beschluss_download.pdf?_blob=publicationFile&v=2, zuletzt abgerufen am 10.02.2022.

Ausnahme bildete die Berechnung der Eigenkapitalverzinsung, da diese in den handelsrechtlichen Daten nicht ausgewiesen wird. Dort werden nur ausschüttungsfähige Gewinne („der Jahresüberschuss“) gezeigt.

Für den Malmquist-Index wurden dagegen die kalkulatorischen Daten aus den Effizienzvergleichen herangezogen. Neben den Gesamtkosten (TOTEX) erfolgte auch eine Berechnung auf Grundlage der standardisierten Gesamtkosten (sTOTEX), um eine Vergleichbarkeit der Kapitalkosten zu gewährleisten.

Somit wurde in der Methodik zum einen die handelsrechtliche Welt abgebildet, zum anderen die kalkulatorische. Die verfügbare Datenbasis wurde so umfassend genutzt, nachdem sowohl von der Branche als auch von der Bundesnetzagentur die Nutzung nicht netzbetreiberspezifischer Daten oftmals verworfen bzw. als nicht adäquat angesehen wurde. So konnte eine möglichst effiziente Datenerhebung mit verschiedenen, in der Wissenschaft und Regulierungspraxis anerkannten Berechnungsmethoden verknüpft werden. Andere, greifbar überlegene Methoden, insbesondere Berechnungen auf Basis volkswirtschaftlicher Indexreihen, sind in diesem Zusammenhang nicht zu erkennen.

2.3.2 Betrachtungszeitraum

Die Bundesnetzagentur hat bei den Berechnungen sowohl zum Törnqvist- als auch zum Malmquist-Index die jeweils längstmöglichen Betrachtungszeiträume herangezogen. Längstmöglich knüpft in diesem Fall an die Datenverfügbarkeit an. Wie oben beschrieben wurde bewusst eine möglichst breite Datenbasis gewählt, so dass für die Berechnung des Törnqvist-Index jeweils eine eigene Datenabfrage für die Bereiche Gas und Strom stattfand. Diese erstreckten sich auf die Jahre 2006 bis 2016 bzw. 2017. Das Jahr 2006 wurde gewählt, weil von den Netzbetreibern ein Jahresabschluss erstmalig für das Jahr 2006 vorzulegen war. Für das Jahr 2005 hätten somit kaum belastbare Daten vorgelegen bzw. die Erhebung derselben hätte zu einem hohen Aufwand für die Unternehmen geführt. Für den Malmquist-Index wurden die verfügbaren Daten aus den Effizienzvergleichen herangezogen, zusätzlich hat eine ergänzende Datenerhebung für den Strombereich stattgefunden.

Bei der Beurteilung des Betrachtungszeitraums (oder auch Stützintervalls) weiterhin zu beachten sind folgende Aspekte:

- Da aus den Berechnungen auf Basis des Stützintervalls eine Prognose für die Zukunft abgeleitet werden soll, sollten zunächst die (Rahmen-)Bedingungen im Stützintervall denen des Prognosezeitraums ähnlich sein. Dies könnten die Jahre sein, die näher an der Gegenwart liegen bzw. Jahre sein, in denen möglichst die gleichen äußerlichen Rahmenumstände gelten wie im Prognosezeitraum.
- Auf der anderen Seite ist die Energiewirtschaft ein Sektor mit längerfristigen Investitionszyklen. Wird nur ein (zu) kurzer Zeitabschnitt gewählt, so besteht die Gefahr, dass diese Zyklen nicht sachgerecht berücksichtigt werden. Daher sollte grundsätzlich ein möglichst den Investitionszyklen nahekommender, langer Zeitraum betrachtet werden.

- Bei der Wahl des Stützintervalls ist auch die Datengüte zu berücksichtigen. Die Daten sollten für alle verwendeten Jahre belastbar sein.
- Sehr lange Zeitreihen (d. h. über Jahrzehnte) sind auf Unternehmensebene zumeist (und auch in der hier betrachteten dritten Regulierungsperiode) nicht verfügbar. Nur Daten von Statistikbehörden reichen u. U. sehr weit zurück. Diese beinhalten allerdings wiederum die Gefahr von Strukturbrüchen in den Daten z. B. aufgrund geänderter Erhebungsmethoden.

Bei der Auswahl des Stützintervalls für die dritte Regulierungsperiode musste die Bundesnetzagentur diesen zum Teil widerstrebenden Aspekten gerecht werden. Durch die Auswahl des Stützintervalls konnte zum einen von relativ ähnlichen Rahmenbedingungen ausgegangen werden, zum anderen wurde ein möglichst langer Zeitraum auf Basis valider Daten gewählt, um Konjunkturzyklen bestmöglich abzubilden.

2.3.3 Mittelwertbildung

Eine weitere Problematik stellte sich bei der Mittelwertbildung innerhalb der Methoden Törnqvist und Malmquist. Dies betrifft zunächst die Frage, ob bei der Ergebnisaggregation das arithmetische oder das geometrische Mittel anzuwenden ist.

„Das arithmetische Mittel [...] ist gleich der Summe aller Beobachtungen, geteilt durch die Anzahl dieser Beobachtungen. Es dient zur repräsentativ-nivellierenden Mittelung (Informationsverdichtung) für mehrere bis zahlreiche nicht zu heterogene Beobachtungen.“⁷²

Das geometrische Mittel dagegen wird bei relativen Änderungen der Merkmalsausprägungen (z. B. Wachstums-, Zuwachsraten, Produktionssteigerungen) berechnet.⁷³ Das geometrische Mittel ist dabei stets kleiner als oder gleich dem arithmetischen Mittel.⁷⁴

Auch die Reihenfolge der Aggregation kann eine Rolle bei der Mittelwertbildung spielen. Setzt sich ein für eine bestimmte Periode berechneter Wert aus unterschiedlichen Einzelindizes zusammen, ergeben sich zwei Möglichkeiten, wie Abbildung 2-7 zeigt.

⁷² Sachs und Hedderich (2009), S. 74.

⁷³ Sachs und Hedderich (2009), S. 82.

⁷⁴ Für eine ausführlichere Diskussion mit einer Beispielrechnung siehe Niederprüm, Stronzik und Wissner (2020), S. 90 ff. und Liebe et al. (2017), S. 40 ff.

Abbildung 2-7: Mittelwertbildung von zusammengesetzten Indexreihen

Jahr	1	2	3	
Output	x	y	z	
Input	a	b	c	
TFP (1)				TFP (2)



Quelle: Eigene Darstellung

Die Aggregation der Veränderungsraten (x , y , z bzw. a , b , c) kann, wie das Beispiel für die Berechnung der Totalen Faktorproduktivität (TFP) zeigt, sowohl in horizontaler als auch in vertikaler Weise erfolgen. Bei der horizontalen Vorgehensweise wird für jedes Jahr ein TFP-Wert gebildet (im Beispiel TFP (1); Output dividiert durch Input) und die so errechneten Werte mathematisch gemittelt. Im Beispiel ist dies also ein Mittelwert aus der Reihe: x/a ; y/b ; z/c .

Bei der vertikalen Vorgehensweise (TFP 2) werden die einzelnen Indexreihen (in diesem Fall Output bzw. Input) über die Jahre gemittelt (x ; y ; z bzw. a ; b , c) und anschließend dividiert. Es lässt sich zeigen, dass bei der Wahl des geometrischen Mittels kein Unterschied im Ergebnis besteht. Bei der Wahl des arithmetischen Mittels hingegen können die Werte abweichen. Bei der grundsätzlichen Wahl des geometrischen Mittels spielt die Reihenfolge der Mittelwertbildung also keine Rolle.

Im Bereich des Malmquist-Indexes stellt sich die Frage, wie die Ergebnisse der verschiedenen Methoden (SFA und DEA) aggregiert werden. Auch hier wurden im Laufe der Verfahren der dritten Regulierungsperiode verschiedene Vorschläge gemacht.

Die Mittelwertbildung beim Malmquist-Index zur Verdichtung der Einzelergebnisse folgt einer konsistenten Logik.⁷⁵ Analog dem Effizienzvergleich werden vier Grundspezifikationen gerechnet (SFA und DEA jeweils mit standardisierten und nicht standardisierten Kosten). Die Ermittlung des Frontier-Shifts erfolgt dabei jeweils auf Basis der Einbeziehung zweier benachbarter Perioden t und $t+1$. Gemäß der Logik des Malmquist-Indexes kann der Frontier-Shift zum einen auf Basis der Vergleichsparameter der ersten Periode t oder auf Basis der Periode $t+1$ ermittelt werden. Beim ersten Fall dienen die Outputparameter von Periode t als Referenztechnologie und beim zweiten Fall die der Periode $t+1$.

Es werden mithin acht Spezifikationen je Frontier-Shift gerechnet. Diese müssen nun zu einer Zahl für die Produktivitätsänderung zwischen t und $t+1$ verdichtet werden. Die Mittelwertbildung hin zum Xgen erfolgt vierstufig:

⁷⁵ Vgl. nachfolgend Liebe et al (2017), S. 64ff.

1. Für jede der acht Spezifikationen: Der Malmquist-Index für eine Spezifikation weist unternehmensindividuelle Werte für den Frontier-Shift aus. Diese werden mittels eines **ungewichteten arithmetischen Mittels** zu einer Zahl je Spezifikation zusammengefasst, da es sich um eine Aggregation über den Querschnitt aller Netzbetreiber handelt und alle Netzbetreiber mit dem gleichen Gewicht in die Betrachtungen eingehen sollen. Der Beitrag eines Unternehmens zum technischen Fortschritt des Sektors ist unabhängig von der Größe eines Unternehmens. Eine Größengewichtung wäre nur dann sachgerecht, wenn nachgewiesen werden könnte, dass der Frontier-Shift mit der Unternehmensgröße korreliert.
2. Für gleiche Spezifikationen unterschiedlicher Vergleichsperioden: Um die beiden Periodenwerte einer Grundspezifikation (DEA oder SFA jeweils mit TOTEX oder sTOTEX) zu aggregieren, wird analog dem Vorgehen beim Malmquist Index das **geometrische Mittel** verwendet. Der Malmquist-Index verwendet diese Form, um die beiden Ergebnisse für die Vergleichsparameter der beiden Perioden zu einer Größe zu verdichten, da ex ante nicht bestimmt werden kann, welche der beiden Ansätze überlegen ist.⁷⁶
3. In einem dritten Schritt werden nun die vier Grundspezifikationen mittels eines **ungewichteten arithmetischen Mittels** zu einer Zahl je Frontier-Shift verdichtet. Dies folgt der Überlegung, dass alle vier Ansätze der Grundspezifikation jeweils ihre spezifischen Vor- und Nachteile aufweisen, wobei keine eindeutige Vorteilhaftigkeit einer Spezifikation bestimmbar ist.⁷⁷ Diese Argumentation wird auch vom Bundesgerichtshof unterstützt.⁷⁸ Entgegen der Ermittlung der individuellen Effizienzwerte nach § 12 Abs. 3 und 4a ARegV bestehe beim Xgen kein besonderes Schutzbedürfnis, da keine individuellen Effizienzvorgaben ermittelt werden, sondern der technische Fortschritt für den gesamten Sektor.
4. Die so verdichteten Ergebnisse der verschiedenen Frontier-Shifts bilden eine Zeitreihe, wie sich der Frontier-Shift über die Zeit entwickelt hat. Daher werden – analog dem Vorgehen beim Törnqvist-Index – die Ergebnisse mittels des **geometrischen Mittels** zu einer Endgröße zusammengefasst.

2.3.4 Validierungsverfahren

Die Ergebnisse der Berechnung des generellen sektoralen Produktivitätsfortschritts können auf unterschiedliche Weise validiert werden. Es existiert daher nicht nur eine einzige (richtige) Validierungsmethode.

Zunächst ist festzustellen, dass durch die Wahl zweier Berechnungsmethoden (Törnqvist- und Malmquist-Index) bereits die Möglichkeit genutzt wurde, die Ergebnisse entsprechend einzuordnen. Die Bundesnetzagentur hat sich im Rahmen der Festlegungen für die dritte Regulierungsperiode jeweils für den niedrigeren der beiden berechneten

⁷⁶ Vgl. z. B. Bogetoft und Otto (2011), S. 43.

⁷⁷ Vgl. z. B. Liebe et al. (2017), S. 55.

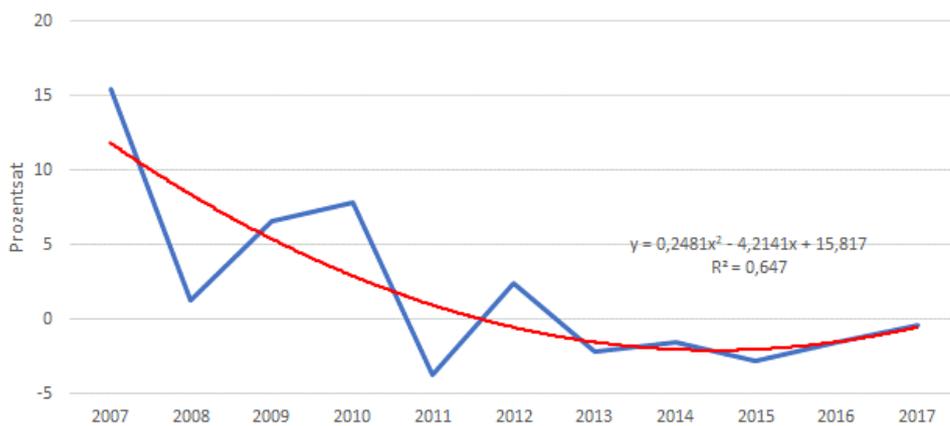
⁷⁸ Vgl. Bundesgerichtshof (2020), insbesondere Rn. 121.

Werte entschieden und somit in der Abwägung zwischen den Interessen der Netzbetreiber auf der einen und der Netznutzer auf der anderen Seite zu Gunsten der Netzbetreiber entschieden. Auch wurde sowohl mit kalkulatorischen als auch mit handelsrechtlichen Daten gerechnet, was ebenfalls zur Plausibilität und Robustheit der Ergebnisse beiträgt.

Des Weiteren wurden während der Verfahren verschiedene Vorschläge für die Validierung der Ergebnisse vorgetragen. Im Folgenden wird auf einzelne dieser Verfahren nochmals eingegangen und ihre Vor- und Nachteile aufgezeigt.

Eine Möglichkeit der Validierung besteht darin, die Entwicklung der betrachteten Jahre grafisch darzustellen und den ablesbaren Trend fortzuführen. In diesem Beispiel erfolgt die Bildung des Xgen jährlich mittels Törnqvist-Index (vgl. dazu auch Abschnitt 4.2.6). In Abbildung 2-8 ist dies für den Stromsektor sichtbar.

Abbildung 2-8: Entwicklung des Xgen im Stromsektor



Die eingezeichnete rote Trendlinie bildet dabei die Entwicklung des gesamten Zeitraumes ab. Würde man sie nach der abgebildeten Formel weiterführen ergäbe sich für die Jahre 2018 bis 2021 ein durchschnittlicher Xgen von 4,45 %.

Der Vorteil dieser Methode ist die Möglichkeit, aufgrund der Vergangenheitswerte eine Prognose für die Zukunft abzugeben. Allerdings muss jeweils die passende Funktion gefunden werden. In diesem Beispiel verfügt eine polynomische Funktion über den höchsten Erklärungsgehalt (höchstes R^2). Allerdings sind auch andere Funktionen denkbar, die zu anderen Ergebnissen führen würden. Dieser Ansatz basiert somit auf der Annahme, dass der für das Stützintervall gefundene funktionale Zusammenhang auch in der Zukunft gültig ist. Insbesondere bei nichtlinearen Zusammenhängen wie in der obigen Abbildung ist dies oft nicht unproblematisch, wie das aufgeführte Ergebnis für den Xgen für die nächste Regulierungsperiode verdeutlicht. Obwohl also nach der gewählten Funktion der Xgen für die nächsten Jahre (wieder) positiv zu werden scheint, kann dies nur ein Anhaltspunkt sein.

Eine weitere Möglichkeit zur Validierung besteht in der Berechnung der Prognosegüte. Dabei wird das Intervall vorhandener Beobachtungen in ein Stütz- und ein Prognoseintervall unterteilt, wobei alle Beobachtungen in der Vergangenheit liegen. Für das Stützintervall wird analog Abbildung 2-8 ein funktionaler Zusammenhang für den Verlauf des Xgen über die Zeit geschätzt. Die Ergebnisse dieser Schätzung dienen dann zur Bestimmung der Prognosegüte. Dabei wird evaluiert, wie gut die Schätzung für das Stützintervall auch für das Prognoseintervall passt bzw. wie hoch die Schätzfehler sind. Je höher diese Fehler sind, desto geringer ist tendenziell die Prognosegüte. Somit kann festgestellt werden, wie gut bzw. genau die Vorhersage war. Diese Methode besitzt den Vorteil, dass auf tatsächlich vorhandenen Daten aufgebaut werden kann.

Auf der anderen Seite müssen auch für diese Methode jährliche Berechnungen des Xgen erfolgen. Je kürzer die für die Prognose verwendeten Zeitreihen sind, desto eher ist die Methode anfällig für Ausreißer. Diese Anfälligkeit ist höher als im vorhergehenden Beispiel, da der vorhandene Zeitraum noch einmal in zwei Zeitintervalle unterteilt wird, so dass sich ceteris paribus die Beobachtungsanzahl für die Ableitung des funktionalen Zusammenhanges für Xgen noch einmal reduziert. Hierfür werden nur die Beobachtungen des Stützintervalls verwendet. Es ist vielmehr zu erwarten, dass ein möglichst langer Betrachtungszeitraum die besten Ergebnisse liefert.

Als Validierungsverfahren kann bei Stichprobenerhebungen auch ein Signifikanztest eingesetzt werden, der Aufschluss gibt, ob ein Ereignis mit einer gewissen Wahrscheinlichkeit nur zufällig aufgetreten ist. Ob eine solche Vorgehensweise auch für Vollerhebungen, wie bei der vorliegenden Berechnung des Xgen, sinnvoll ist, ist wissenschaftlich zumindest umstritten.⁷⁹ Manche Lehrbücher sind hier auch klar in ihrer Position:⁸⁰

„Bei einer Voll- oder Totalerhebung werden alle Mitglieder der Grundgesamtheit in die Erhebung miteinbezogen. Dies ist nur bei überschaubaren und bekannten zugänglichen Grundgesamtheiten möglich (z. B. bei kleineren Bevölkerungs- und Personengruppen, Mitarbeitenden kleinerer Firmen oder Institutionen). Nur wenn man alle kennt und auch ansprechen kann bzw. im Zugriff hat, kann man sie auch analysieren (etwa Zeitungsartikel bei einer Inhaltsanalyse). Wenn man wirklich ALLE Teile der Grundgesamtheit erreicht, erzielt eine Vollerhebung immer genauere Resultate als eine Stichprobe. Man braucht sich keinerlei Gedanken über Repräsentativität oder eine Hochrechnung von Ergebnissen zu machen: Die Ergebnisse der Vollerhebung stellen ja ein vollkommenes Abbild der Grundgesamtheit dar. Wenn man die Grundgesamtheit vollzählig in die Erhebung miteinbeziehen kann, gibt es weder Schwankungsbreiten der Ergebnisse (Stichprobenfehler) noch statistische Signifikanztests, worauf man Rücksicht nehmen müsste [...] Derartige Berechnungen hängen immer mit dem Übertragen von Stichprobenergebnissen auf Grundgesamtheiten zusammen. Wenn aber bereits „alles“ erhoben wurde, muss nichts mehr übertragen werden. Die Grundgesamtheit wurde bereits vollkommen und ohne Stichprobenziehung erfasst. Die Ergebnisse treffen also ohne jegliche Einschrän-

⁷⁹ Vgl. z. B. Behnke (2005).

⁸⁰ Braunecker (2016), S. 39

kung zu. [...] Eine Stichprobe ist immer dann gegenüber einer Vollerhebung zu bevorzugen, wenn man die Grundgesamtheit nicht kennt, wegen ihrer Größe nicht erforschen oder in ihrer Gesamtheit nicht erreichen kann.“

Signifikanztests können aber im Einzelfall sinnvoll sein, falls man den erhobenen Daten nicht vollständig vertraut (z. B. aufgrund von Messfehlern) oder die Grundgesamtheit nicht vollständig erreicht wird.

2.4 Zwischenfazit

Wie gezeigt existieren also unterschiedliche Möglichkeiten, Ergebnisse zu validieren. Konkrete Validierungen zum Malmquist-Index der vierten Regulierungsperiode sind in Abschnitt 5 aufgeführt, zum Törnqvist-Index finden sich diese im Anhang. Dort werden die Ergebnisse von Sensitivitätsrechnungen für die Variation des Outputparameters und unterschiedlicher Mittelwertbildungen aufgeführt. Ebenfalls erfolgt eine Prognose aus dem Trend der vorliegenden Jahre sowie eine Betrachtung verschiedener Zeiträume für das Stützintervall. Es zeigt sich, dass die Festlegung für die dritte Regulierungsperiode im Gasbereich als sehr konservativ einzuschätzen ist, da die tatsächliche Entwicklung des Xgen weit höher verlief. Wie auch konkret in Abschnitt 3 für einzelne Bestandteile des Xgen diskutiert, erfolgte die Festlegung des Xgen nach dem Stand der Wissenschaft. Eine greifbar überlegene Methodik ist nicht erkennbar.

3 Ermittlung der gesamtwirtschaftlichen Bestandteile nach § 9 ARegV

Die gesamtwirtschaftlichen Bestandteile des Xgen nach § 9 ARegV werden residual über die Änderungsrate des Verbraucherpreisindex (VPI) ermittelt. Dies leitet sich aus folgenden Überlegungen ab:⁸¹

Im Wettbewerb gilt, dass die Änderung der Outputpreise ($\Delta P_{Output,t}$) der Änderung der Inputpreise ($\Delta P_{Input,t}$) abzüglich des technischen Fortschritts (ΔTF_t) entspricht:

$$(3-1) \quad \Delta P_{Output,t} = \Delta P_{Input,t} - \Delta TF_t.$$

In der Gesamtwirtschaft (GW) entspricht die Änderung der Outputpreise der Inflation, ausgedrückt als Änderung des Verbraucherpreisindex:

$$(3-2) \quad \Delta P_{Output,t}^{GW} = \Delta VPI_t.$$

Werden diese Zusammenhänge für den Xgen im deutschen Kontext verwendet, folgt

$$(3-3) \quad X_{Gen,t} = (\Delta TF_t^{Netz} - \Delta TF_t^{GW}) + (\Delta P_{Input,t}^{GW} - \Delta P_{Input,t}^{Netz}) \\ = (\Delta TF_t^{Netz} - \Delta P_{Input,t}^{Netz}) + \Delta VPI_t.$$

Diese Vereinfachung kann unter der Annahme vorgenommen werden, dass die Gesamtwirtschaft in Deutschland wettbewerbsmäßig organisiert ist. Die Vereinfachung dient vor allem der Vermeidung von Fehlerquellen, die unter Umständen bei der separaten Ermittlung der beiden Einzelbestandteile auftreten, der Änderungsrate der gesamtwirtschaftlichen Einstandspreise sowie die Rate des technischen Fortschritts für die Gesamtwirtschaft. Während die Ermittlung der Änderungsrate der totalen Faktorproduktivität auf gesamtwirtschaftlicher Ebene noch relativ unumstritten ist, gibt es keinen deutschlandweiten Indikator für die Abbildung der Inputpreise.

Die Anwendung der Residualmethode wurde im Rahmen der Gerichtsverfahren zu den Festlegungen des Xgen für die dritte Regulierungsperiode kritisiert. Insbesondere entsprechen das gewählte Vorgehen nicht dem Stand der Wissenschaft und verstoße gegen § 9 Abs. 1 ARegV, weswegen ein sogenannter Differenzenansatz verwendet werden sollte, der die beiden Bestandteile separat ermittele. Ein weiterer Kritikpunkt betrifft die Annahme vollkommener Konkurrenz für die deutsche Volkswirtschaft.⁸²

Sowohl der Bundesgerichtshof als auch das Oberlandesgericht Düsseldorf haben jedoch die Residualmethode als nicht zu beanstandende Vorgehensweise eingestuft.⁸³ So sei das gewählte Vorgehen mit § 9 Abs. 1 und 3 S. 1 ARegV vereinbar und rekurriere auf wissenschaftlich anerkannte Zusammenhänge. Eine systematische Verzerrungswirkung sei nicht feststellbar. Der vom Oberlandesgericht Düsseldorf beauftragte Sachverständige spricht im Zusammenhang mit dem von der Branche geforderten Differenzenansatz

⁸¹ Vgl. nachfolgend Liebe et al. (2017), S. 44ff.

⁸² Vgl. z. B. Oberlandesgericht Düsseldorf (2022), insbesondere S. 15ff.

⁸³ Siehe Bundesgerichtshof (2021), insbesondere Rn. 39 sowie Oberlandesgericht Düsseldorf (2022), insbesondere S. 33.

von einer „vermeintlichen“ Kompensationswirkung in Hinblick auf mögliche Fehlerquellen.⁸⁴ Da eine Überlegenheit des Differenzenansatzes nicht vorhanden sei, sieht der Bundesgerichtshof den Verzicht auf eine Plausibilisierung der Ergebnisse der Residualmethode durch den Differenzenansatz zudem als zulässig an.⁸⁵

In Hinblick auf die Annahme der Wettbewerbsfähigkeit für die Gesamtwirtschaft urteilt der Bundesgerichtshof, dass perfekter Wettbewerb nicht erforderlich sei. Vielmehr sei es hinreichend, wenn die Wettbewerbsintensität und die allokativen Ineffizienzen in der Gesamtwirtschaft unverändert geblieben seien.⁸⁶

Obwohl die Residualmethode seitens der Gerichte uneingeschränkt bestätigt wurde, werden die zwei wesentlichen inhaltlichen Kritikpunkte der Branche im Rahmen dieses Gutachtens noch einmal aufgenommen und diskutiert. In Abschnitt 3.1 wird der Stand der Wissenschaft anhand der Literatur und der internationalen Regulierungspraxis erörtert. Anschließend wird die Annahme der Wettbewerbsfähigkeit für die deutsche Volkswirtschaft auf Basis gängiger Wettbewerbsindikatoren kritisch hinterfragt. Abschließend wird ein kurzes Zwischenfazit gezogen.

3.1 Stand der Wissenschaft

3.1.1 Wissenschaftliche Literatur

Die Annahme vollkommener Konkurrenz und daraus folgender gesamtwirtschaftlicher Zusammenhänge, wie sie im Rahmen der Residualbetrachtung erfolgt, ist Bestandteil vieler wissenschaftlicher Abhandlungen und somit Stand der Wissenschaft. So geben z. B. Jafari und Britz (2018) einen Überblick über den aktuellen Stand der Forschung bei rechenbaren Allgemeinen Gleichgewichtsmodellen zur Abbildung gesamtwirtschaftlicher Zusammenhänge. Im Kontext dieser Computable General Equilibrium (CGE) Models ist vollkommener Wettbewerb die Grundvariante für die Abbildung der meisten Sektoren. Nur in ausgewählten Sektoren, auf denen in der Regel der Fokus der Betrachtung liegt, wird unvollkommener Wettbewerb modelliert. Ein ähnliches Bild zeichnen Robson et al. (2018), die einen Überblick über CGE-Modelle im Transportsektor geben.

Die Annahme vollkommener Konkurrenz bildet zudem die Grundlage für die Ableitung der Formel für den Xgen in Bernstein und Sappington (1999). In Formel (2.6) weisen die Autoren explizit auf den Zusammenhang hin, dass sich bei vollkommener Konkurrenz die Änderung der Outputpreise aus der Differenz der Änderung der Inputpreise und der Rate des technischen Fortschritts ergibt. Diese Formel bildet dabei den Ausgangspunkt für die weiteren Ausführungen der Autoren, in denen sie letztendlich die bekannte Formel für den Xgen ableiten, der sich aus einem Produktivitäts- und einem Inputpreisdifferenzial zwischen reguliertem Sektor und der Gesamtwirtschaft ergibt.⁸⁷ Allerdings weisen die

⁸⁴ Vgl. von Auer (2019), S. 54.

⁸⁵ Vgl. Bundesgerichtshof (2021), insbesondere Rn. 44.

⁸⁶ Vgl. Bundesgerichtshof (2021), insbesondere Rn. 43.

⁸⁷ Siehe Formel (2.11) in Bernstein und Sappington (1999).

Autoren auch daraufhin, dass die Differenzialbetrachtung nur erforderlich ist, wenn die Kostenbasis des regulierten Unternehmens mit einem gesamtwirtschaftlichen Outputpreisindex inflationiert wird. Bernstein und Sappington (1999, S. 9) führen dazu aus:⁸⁸

„Wenn also die Preise eines regulierten Unternehmens anfänglich so festgelegt werden, dass kein (außergewöhnlicher) Gewinn erzielt wird, und wenn die Preise des Unternehmens anschließend um eine Rate geändert werden müssen, die der Differenz zwischen der Rate, mit der seine Inputpreise steigen, und seiner Produktivitätssteigerung entspricht, dann wird das regulierte Unternehmen weiterhin keinen extranormalen Gewinn erzielen, genau wie auf einem wettbewerbsorientierten Markt.“

Daraus folgt ferner, dass, wenn die Kostenbasis eines regulierten Unternehmens mit einem sektoralen Inputpreisindex inflationiert wird, der Xgen allein der Rate des sektoralen technischen Fortschritts entspricht. Eine Differenzialbetrachtung ist daher auch nach Bernstein und Sappington obsolet, wenn die Kostenbasis des regulierten Unternehmens mit einem sektoralen Inputpreisindex inflationiert wird.⁸⁹

3.1.2 Internationale Regulierungspraxis

Nachfolgend werden mit den Niederlanden, Österreich und dem Vereinigten Königreich drei Beispiele aus der internationalen Regulierungspraxis betrachtet, die ebenfalls einen generellen X-Faktor im Rahmen der Regulierung der Energienetze anwenden. Es wird untersucht, inwiefern dort ähnliche Zusammenhänge für den Xgen verwendet werden, wie dies in Deutschland durch die Residualbetrachtung erfolgt.

3.1.2.1 Niederlande

In den Niederlanden erfolgt eine analoge Handhabung von Strom- und Gasnetzen. Die Kostenbasis wird während einer Regulierungsperiode mittels der Differenz aus Verbraucherpreisindex und einem X-Faktor inflationiert. Zur Ermittlung des X-Faktors wird ein Yardstick-Ansatz verwendet, wobei die Preise bzw. Erlöse durch einen Bezug zu den Durchschnittskosten des Sektors festgelegt werden. Dabei werden die Erlöse je Output-einheit betrachtet. Bei Strom werden die Outputparameter Anschlussleistung (in kW), die Menge der verteilten Strommenge (in kWh) und die Anzahl der Kunden betrachtet. Bei Gas sind die Kapazität der Anschlüsse (in Kubikmeter Gas pro Stunde und Jahr) und die Anzahl der Kunden die relevanten Outputs.⁹⁰

Die Ermittlung des X-Faktors in den Niederlanden folgt folgender in vereinfachter Form dargestellten Systematik:⁹¹

⁸⁸ Eigene Übersetzung.

⁸⁹ Dieser Aspekt hat zwar für Deutschland keine unmittelbare Relevanz, da die Kostenbasis mit dem VPI inflationiert wird. Allerdings ist er wichtig für die Ableitung möglicher Kongruenzen bei der Betrachtung von Beispielen der internationalen Regulierungspraxis in Abschnitt 3.1.2.

⁹⁰ Vgl. nachfolgend ACM (2021a) für Strom (insbesondere S. 30ff.), ACM (2021b) für Gas (insbesondere S. 27ff.) sowie ACM (2021c) für die formale Darstellung des X-Faktors.

⁹¹ Siehe auch CEER (2020a: 20ff.).

1. In einem ersten Schritt werden für jeden Netzbetreiber auf Basis der tatsächlichen Outputmengen und der Tarife die tatsächlichen Erlöse im Basisjahr ermittelt.
2. Im zweiten Schritt werden für jeden Netzbetreiber die erwarteten effizienten Kosten für das letzte Jahr der kommenden Regulierungsperiode (Zieljahr) berechnet. Dies erfolgt wiederum in fünf Unterschritten:
 - a. Zunächst werden für jeden Netzbetreiber die Kosten für das Basisjahr berechnet, wobei der WACC der kommenden Regulierungsperiode zugrunde gelegt wird.
 - b. Im nächsten Unterschritt werden die durchschnittlichen Tarife (Erlöse je Output-einheit) aller Netzbetreiber im Basisjahr für die betrachteten Outputs ermittelt. Das Mengengerüst wird dabei grundsätzlich aus dem Basisjahr auch ins Zieljahr fortgeschrieben.
 - c. Auf Basis der durchschnittlichen Tarife werden unter Zugrundlegung des genannten Mengengerüsts für die Outputs die erwarteten Erlöse je Netzbetreiber für das Zieljahr bestimmt. Es wird somit angenommen, dass alle Netzbetreiber die gleichen Erlöse je Output-einheit realisieren, wobei als Nebenbedingung gilt:

$$\sum_i \text{Erlös}(\text{Schritt1})_i = \sum_i \text{Erlös}(\text{Schritt2c})_i$$
 - d. Im vierten Unterschritt werden die effizienten Kosten pro Output-einheit für das Zieljahr berechnet mittels $\sum_i \text{Kosten}(\text{Schritt2a})_i / \sum_i \text{Erlös}(\text{Schritt2c})_i$. Diese effizienten Kosten je Output-einheit sind für alle Netzbetreiber gleich. Im Endeffekt handelt es sich dabei um den durchschnittlich insgesamt vom Sektor zu erbringenden Produktivitätsfortschritt über die kommende Regulierungsperiode.
 - e. Im letzten Unterschritt werden auf Basis dieser effizienten Kosten je Output-einheit aus Unterschritt 2d unter Zugrundelegung des genannten Mengengerüsts für die Outputs für jeden Netzbetreiber die effizienten Kosten für das Zieljahr bestimmt, indem die effizienten Kosten je Output-einheit mit dem Mengengerüst für den Netzbetreiber multipliziert werden.
3. Abschließend wird der X-Faktor für jeden Netzbetreiber ermittelt, der sich aus dem linearen Anpassungspfad zwischen seinen tatsächlichen Erlösen im Basisjahr (Schritt 1) und den effizienten Kosten des Zieljahres (Schritt 2e) ergibt. Diese X-Faktoren sind für jeden Netzbetreiber unterschiedlich und setzen sich aufgrund der Berechnungslogik aus einer individuellen Komponente und einer Sektorkomponente zusammen.

Wie in Deutschland wird in den Niederlanden die Kostenbasis mit dem VPI inflationiert und es kommt ein X-Faktor zur Anwendung. Allerdings ist der Ansatz bei näherer Betrachtung aufgrund von Unterschieden in der Regulierungssystematik nicht unmittelbar mit dem deutschen Ansatz vergleichbar. So kommt in den Niederlanden ein Yardstick-Ansatz (Bezug auf die durchschnittlichen Kosten des Sektors) zur Anwendung und die Kosten und Erlöse werden zur Ermittlung der effizienten Kosten auf die Output-einheit bezogen. Dadurch beinhaltet der X-Faktor sowohl eine netzbetreiberspezifische (analog dem individuellen X-Faktor in Deutschland) als auch eine sektorweite Komponente (analog dem Xgen in Deutschland). Das CEER spricht in diesem Kontext daher auch davon,

dass der X-Faktor in den Niederlanden eher ein Preisdifferenzial als eine Produktivitätsvorgabe darstellt.⁹²

3.1.2.2 Österreich

Auch in Österreich ist das Vorgehen bei Strom und Gas hinsichtlich des vorliegenden Kontextes ähnlich.⁹³ Österreich hat eine Erlösobergrenzenregulierung, wobei die Kostenbasis mittels eines sektoralen Inputpreisindex, dem sogenannten Netzbetreiberpreisindex, inflationiert wird. Analog dem deutschen System werden ein individueller und genereller X-Faktor ermittelt. Der individuelle X-Faktor wird mittels eines Benchmarkings bestimmt. Der generelle X-Faktor wird direkt als Änderung der sektoralen Faktorproduktivität berechnet. Wurde letzterer früher mittels eines Törnqvist-Ansatzes aus Daten der volkswirtschaftlichen Gesamtrechnung ermittelt, wird neuerdings auf ökonometrische Verfahren auf Basis von Netzbetreiberdaten zurückgegriffen.⁹⁴ Kernpunkt ist ein Ansatz mittels einer OLS mit Zeittrend.

Eine Besonderheit besteht in Österreich darin, dass der Xgen allein auf die OPEX wirkt. Die Kapitalkosten werden über einen Kapitalkostenabgleich nicht in diese Systematik einbezogen. Daher wird die sektorale Faktorproduktivität als partielle Faktorproduktivität berechnet. Mithin basieren die Schätzungen nur auf den OPEX. Für mögliche OPEX/CAPEX-Substitutionsbeziehungen wird jedoch im Rahmen von Sensitivitätsbetrachtungen kontrolliert.

Der Verzicht auf eine Differenzialbetrachtung à la Bernstein und Sappington bestätigt somit indirekt das Vorgehen im deutschen Kontext, da bei einer Inflationierung der Kostenbasis mittels eines sektoralen Inputpreisindexes auch nach Bernstein und Sappington eine Differenzialbetrachtung obsolet ist.⁹⁵ Bei einer Inflationierung der Kostenbasis mit einem sektoralen Inputpreisindex bestimmt sich der Xgen auch nach Bernstein und Sappington somit allein aus der Rate des sektoralen technischen Fortschritts.

3.1.2.3 Vereinigtes Königreich

Seit dem 1.4.2021 ist im Vereinigten Königreich die zweite Periode der RIIO-Regulierung (RIIO-2) in Kraft.⁹⁶ Im Rahmen dieser Regulierung wird ein X-Faktor ermittelt, der sich aus zwei Größen zusammensetzt:

- OE: Ongoing Efficiency Target,
- RPE: Real Price Effects,

wobei gilt

⁹² Vgl. CEER (2020b: 84).

⁹³ Für die aktuelle Regulierungsperiode siehe E-Control (2017) für Gas und E-Control (2018) für Strom.

⁹⁴ Siehe Schmitt et al. (2017) für Gas und Stronzik und Wissner (2018) für Strom.

⁹⁵ Vgl. Abschnitt 3.1.1.

⁹⁶ Vgl. nachfolgend Ofgem (2018, 2020) und CEPA (2020a, 2020b).

$$(3-4) \quad X = OE - RPE.$$

OE bildet dabei den sektoralen technischen Fortschritt ab. Wie in Abschnitt 2.1.3.1 bereits ausgeführt, wurden für die Bestimmung von OE vor allem unterschiedliche Produktivitäten auf Basis von Daten aus der EU-KLEMS Datenbank mittels einer Törnqvist-Methodik berechnet. Es wurden sowohl totale Faktorproduktivitäten (u. a. mittels der Bildung von Vergleichssektoren)⁹⁷ als auch partielle Faktorproduktivitäten (u. a. Arbeitsproduktivität) bestimmt. Das Stützintervall für die Berechnungen ist 1997 bis 2016. Auf Basis des durch diese Produktivitätsberechnungen aufgespannten Wertebereichs hat Ofgem final ein OE von 1,15 % für CAPEX und 1,25 % für OPEX festgelegt.⁹⁸

Der zweite Term, RPE, erfasst die über die gesamtwirtschaftliche Inflationsrate hinausgehende sektorale Inputpreissteigerung. Hierzu werden verschiedene Inputpreisindices herangezogen. Es erfolgt eine jährliche Anpassung während der Regulierungsperiode.⁹⁹

Ferner erfolgt für RIIO-2 ein Aufschlag in Höhe von 0,2 % p.a. auf das festgelegte OE.¹⁰⁰ Aufgrund umfangreicher innovationsfördernder Maßnahmen während der vorhergehenden Regulierungsperiode RIIO-1, die durch die Netznutzer über einen Aufschlag auf die Netzentgelte finanziert wurden, erwartet Ofgem einen zusätzlichen technischen Fortschritt in dieser Höhe.

In letzter Konsequenz ähnelt der X-Faktor im Vereinigten Königreich somit dem Vorgehen in Deutschland. Während OE den sektoralen technischen Fortschritt abbilden soll, erfasst RPE die über die gesamtwirtschaftliche Inflationsrate hinausgehende sektorale Inputpreissteigerung. Es gilt mithin:

$$(3-5) \quad OE \approx \Delta TF_{\text{Netz}} \text{ und}$$

$$RPE = \Delta IP_{\text{Netz}} - \Delta VPI$$

Aus (3-4) und (3-5) folgt:

$$(3-6) \quad X = \Delta TF_{\text{Netz}} - (\Delta IP_{\text{Netz}} - \Delta VPI)$$

$$X = (\Delta TF_{\text{Netz}} - \Delta IP_{\text{Netz}}) + \Delta VPI$$

Formel (3-6) entspricht der im Beschluss der Bundesnetzagentur angewendeten Systematik unter Zugrundelegung der Residualbetrachtung für die Gesamtwirtschaft.¹⁰¹

3.2 Annahme der Wettbewerbllichkeit

In diesem Abschnitt wird die Annahme der Wettbewerbllichkeit für die deutsche Volkswirtschaft auf Basis gängiger Wettbewerbsindikatoren kritisch hinterfragt. Es gibt eine Viel-

⁹⁷ Dies ähnelt dem Ansatz der Bildung eines synthetischen Indexes, wie z. B. in Liebe et al. (2017), S. 72ff. ausgeführt ist.

⁹⁸ Vgl. Ofgem (2020), S. 48.

⁹⁹ Ofgem (2020), S. 68f. listet netzbetreiberspezifische Prognosen für die Jahre bis 2026 auf.

¹⁰⁰ Vgl. Ofgem (2020), S. 213.

¹⁰¹ Vgl. Bundesnetzagentur (2018), S. 17.

zahl von Indikatoren mit sehr unterschiedlichen Intentionen zur Messung von Wettbewerb. Die zwei wesentlichen, gängigsten und am weitesten verbreiteten Indikatoren zu Beantwortung der Frage im vorliegenden Kontext, ob Deutschland insgesamt als wettbewerbsfähig organisierte Volkswirtschaft anzusehen ist und wie sich diese unter Umständen über die Zeit entwickelt hat, sind:¹⁰²

- Der Global Competitiveness Index (GCI) des Weltwirtschaftsforums sowie
- Das World Competitiveness Yearbook (WCY) des IMD World Competitiveness Centers.

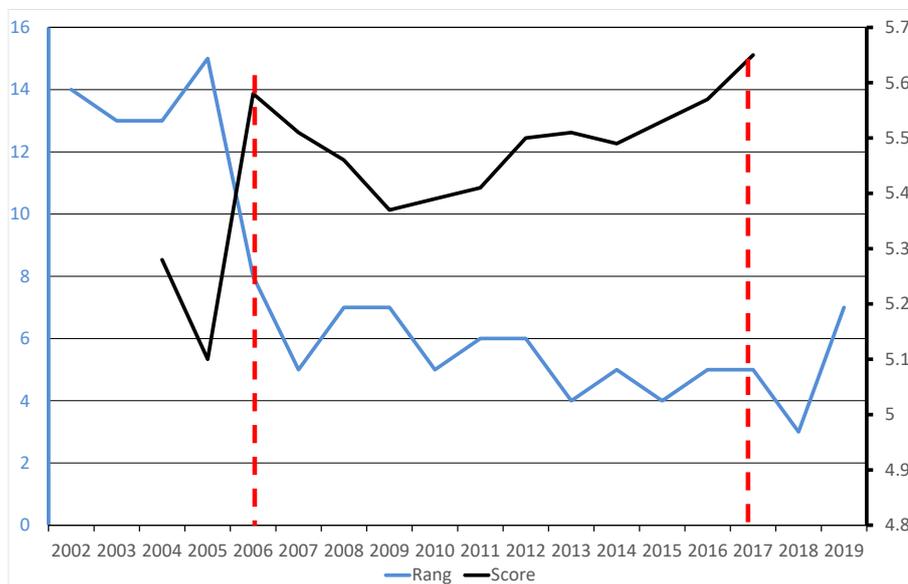
Abbildung 3-1 enthält eine Auswertung der verfügbaren Informationen für den GCI für Deutschland. Der GCI bewertet die Wettbewerbsfähigkeit von rund 140 Ländern auf Basis diverser Einzelindikatoren mittels einer Skala (Score) von 0 (kein funktionsfähiger Wettbewerb) und 6 (vollkommenen funktionsfähiger Wettbewerb). Auf dieser Basis werden die Länder hinsichtlich der in diesen Ländern herrschenden Wettbewerbsfähigkeit in eine Rangordnung gebracht. Verlässliche und konsistente Daten liegen vor allem für den Zeitraum 2006 bis 2017 vor, was durch die rotgestrichelten Linien in Abbildung 3-1 verdeutlicht ist. Vor 2006 sind sowohl die Angaben für den Score als auch für den Rang mit Vorsicht zu interpretieren, da sie auf Sekundärquellen zurückgreifen, die zwar auf das Weltwirtschaftsforum rekurrieren, nicht jedoch direkt von diesem stammen. Ab 2018 hat das Weltwirtschaftsforum seine Bewertungsmetrik für den Score auf eine Skala von 0 bis 100 umgestellt. Aufgrund dieses Strukturbruches sind die Werte für den Score nach 2017 nicht mehr dargestellt. Die Rangangaben stammen jedoch weiterhin vom Weltwirtschaftsforum und sind mit den vorherigen Informationen vergleichbar.

Ab 2006 liegt Deutschland konstant in den TOP10 der betrachteten Länder. Der Score liegt von 2006 und 2017 konstant über 5,4, was auf einen funktionsfähigen Wettbewerb in Deutschland hinweist. Nach einem leichten Rückgang bis 2009 ist der Wert bis 2017 wieder auf sein Niveau von 2006 zurückgekehrt. Die Schwankungsbreite ist somit als gering einzustufen. Die Rangangaben deuten auf eine leichte Verschlechterung der Wettbewerbsfähigkeit in Deutschland nach 2017 hin.¹⁰³ Insgesamt weist der GCI für Deutschland eine hohe Wettbewerbsfähigkeit auf mit nur leichten Schwankungen in der Intensität.

102 Vgl. Sargsyan (2017). Als Beispiel für zunächst dem Namen nach geeignete Indikatoren, die jedoch eine andere Intention verfolgen, seien die Harmonized Competitiveness Indicators (HCI) der Europäischen Zentralbank angeführt. Die Methodik orientiert sich an der des Wechselkurskonzeptes. Die Indikatoren bilden daher eher relative Preis- bzw. Kostenvorteile (z. B. aufgrund vom Lohnniveau) ab, die ein Land auf den Weltmärkten gegenüber potenziellen Wettbewerbern besitzt. Im Fokus stehen somit die Absatzchancen eines Landes, inwiefern es preiswerter in Relation zu Wettbewerbern anbieten kann. Vgl. https://www.ecb.europa.eu/stats/balance_of_payments_and_external/hci/html/index.en.html.

103 Dies wird auch durch die neue Bewertungsmetrik unterlegt. So ist der Score von 2018 auf 2019 von 82,8 auf 81,8 leicht gefallen.

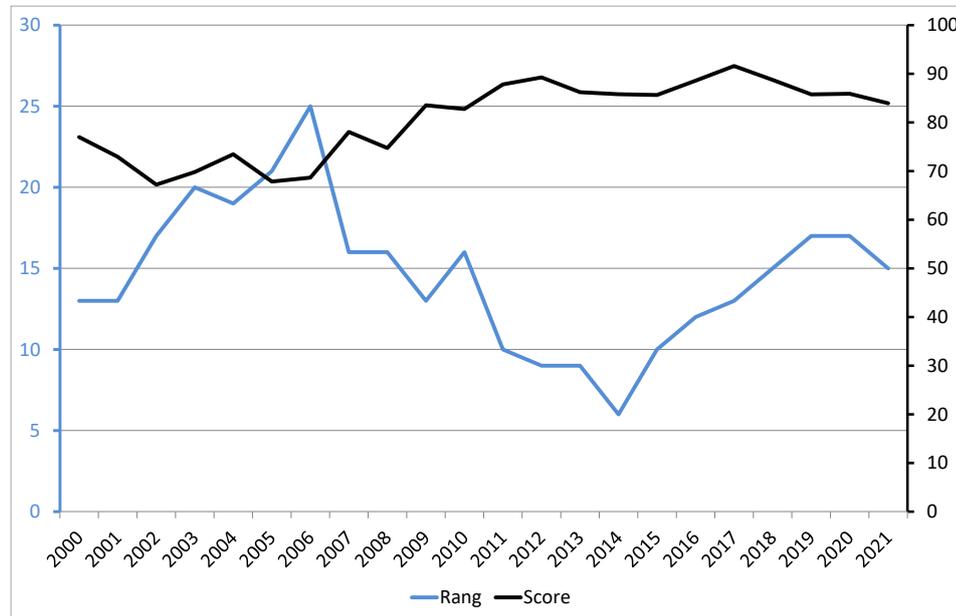
Abbildung 3-1: Entwicklung des Global-Competitiveness-Indexes für Deutschland




Quelle: <https://www.weforum.org/reports/> .

Ein ähnliches Bild zeigt sich bei einer Auswertung der verfügbaren Informationen für den WCY. Die Bewertungssystematik bei WCY ist ähnlich wie beim GCI, allerdings basiert die Bewertungsmetrik durchgehend auf einer Skala von 0 (kein funktionsfähiger Wettbewerb) bis 100 (vollkommenen funktionsfähiger Wettbewerb). Ferner gehen nur rund 70 Länder in die Bewertung mit ein. Verlässliche und konsistente Daten sowohl für den Score als auch für den Rang sind beim WCY für 2000 bis 2021 verfügbar.

Abbildung 3-2: Entwicklung des Indexes World Competitiveness Yearbook für Deutschland



Quelle: <https://worldcompetitiveness.imd.org/customsearch>.

In Abbildung 3-2 sind die Ergebnisse für Deutschland dargestellt für den WCY. Von 2006 bis 2014 weist der Index für Deutschland eine stetige Verbesserung der Wettbewerblichkeit auf. Eine substantielle Verbesserung des Rangs von Position 25 auf Position 6 geht mit einem leichten Anstieg des Scores einher. Seit 2014 hat sich die Rangposition von Deutschland zwar wieder substantiell verschlechtert, wobei der Score jedoch auf einem konstant hohen Niveau verbleibt und nur leichten Schwankungen unterworfen ist. Insgesamt bestätigt der WCY somit die aus dem Verlauf des GCI abgeleiteten Aussagen. Deutschland kann als eine Volkswirtschaft mit einem funktionsfähigen Wettbewerb eingestuft werden, wobei die Intensität nur leichten Schwankungen unterliegt.¹⁰⁴

3.3 Zwischenfazit

Die Anwendbarkeit der Residualmethode zur Darstellung der gesamtwirtschaftlichen Komponenten bei der Bestimmung des generellen sektoralen Produktivitätsfortschritts nach § 9 ARegV wurde in diesem Kapitel einer kritischen Würdigung unterzogen. Es

¹⁰⁴ Gugler und Liebensteiner (2020), S. 11ff ziehen sowohl bzgl. des Vorherrschens eines funktionsfähigen Wettbewerbs als auch bzgl. der Änderung der allokativen Effizienz über die Zeit für Deutschland ein gegenteiliges Fazit. Die hohen Scores sowohl für den GCI als auch für den WCY sind jedoch ein eindeutiges Indiz für einen funktionsfähigen Wettbewerb in Deutschland. Ferner betrachten sie nicht das gesamte Zeitintervall seit 2006 sondern nur einen Teilausschnitt. Dadurch ziehen sie fälschlicherweise das Fazit, dass sich die Wettbewerbsintensität in Deutschland erhöht habe. Die geringen Schwankungen in den Scores deuten vielmehr auf ein normales Rauschen einer relativ konstant verbleibenden Wettbewerblichkeit in Deutschland hin.

konnten keine Anhaltspunkte identifiziert werden, die einer Anwendbarkeit entgegenstehen.

Sowohl der Bundesgerichtshof als auch das Oberlandesgericht Düsseldorf haben bestätigt, dass die Residualmethode sachgerecht ist. Nichtsdestotrotz wurden die zwei wesentlichen Kritikpunkte der Branche an der Residualmethode (kein Stand der Wissenschaft, Annahme vollkommener Konkurrenz nicht sachgerecht für Deutschland) aufgegriffen und noch einmal kritisch geprüft.

Hinsichtlich des Standes der Wissenschaft wurde sowohl die Literatur als auch die internationale Regulierungspraxis betrachtet. Die Annahme vollständiger Konkurrenz ist weiterhin weit verbreitet in ökonomischen Analysen, insbesondere bei der Betrachtung gesamtwirtschaftlicher Zusammenhänge.

Für Rückschlüsse aus der internationalen Regulierungspraxis für den deutschen Ansatz ist jeweils die Spezifität der verwendeten Ansätze zu beachten. Während die Niederlande aufgrund des Yardstick-Ansatzes und des Bezugs auf den Output in Bezug auf die Residualmethode nicht unmittelbar mit dem deutschen Ansatz vergleichbar sind, bestätigen Österreich und das Vereinigte Königreich tendenziell das Vorgehen in Deutschland. In Österreich erfolgt dies indirekt über den Verzicht auf die Differenzialbetrachtung bei einer Inflationierung der Kostenbasis mittels eines sektoralen Inputpreisindex. Wie gezeigt werden konnte, folgt auch das Vereinigte Königreich durch die spezifische Ausgestaltung des X-Faktors einer ähnlichen Logik wie die Bundesnetzagentur.

Eine Auswertung wesentlicher Wettbewerbsindikatoren bestätigt zudem die hohe Wettbewerbsfähigkeit der deutschen Volkswirtschaft, wobei die Intensität nur leichten Schwankungen unterworfen ist. Die Indikatoren deuten somit weder darauf hin, dass die Annahme vollkommener Konkurrenz für Deutschland nicht sachgerecht sei, noch dass es signifikante Änderungen der allokativen Effizienz über den relevanten Zeitraum seit 2006 gab.

4 Törnqvist-Methode

In diesem Abschnitt wird die Törnqvist-Methode in ihrer vorliegenden Anwendung einer wissenschaftlichen Überprüfung unterzogen. Dazu werden die zugrundeliegenden Daten und verschiedene Einzelelemente geprüft und auf eine Veränderung beziehungsweise Fortentwicklung untersucht.

Die wissenschaftliche Bewertung betrifft zwei Aspekte. Zum einen wird allgemein die Frage untersucht, inwiefern Daten sachgemäß sind, um den Sachverhalt bestmöglich abzubilden. Zum anderen wird bezüglich der Methodik auf einzelne Elemente eingegangen und diese werden einer grundlegenden Prüfung unterzogen. Dabei wird dann wiederum, wo es möglich ist, auf die Auswahl verschiedener alternativer Datenquellen eingegangen.

4.1 Daten

Die hier untersuchte Fragestellung beinhaltet eine Darstellung, wie Datenqualität bewertet werden kann, welche Daten für den vorgesehenen Zweck vorhanden sind und wie gut sie sich für eine Berechnung nach der Törnqvist-Methode eignen.

Grundsätzlich können Daten nach verschiedenen Qualitätsmerkmalen beurteilt werden, wie in Abbildung 4-1 ersichtlich ist.

Abbildung 4-1: Qualitätsmerkmale von Daten

Merkmalsklasse	Qualitätsmerkmal
Intrinsische Datenqualität	Glaubhaftigkeit (<i>believability</i>)
	Genauigkeit (<i>accuracy</i>)
	Objektivität (<i>objectivity</i>)
	Reputation (<i>reputation</i>)
Kontextuelle Datenqualität	Mehrwert (<i>value-added</i>)
	Relevanz (<i>relevancy</i>)
	Zeitnähe (<i>timeliness</i>)
	Vollständigkeit (<i>completeness</i>)
	Datenmenge (<i>amount of data</i>)
Repräsentationelle Datenqualität	Interpretierbarkeit (<i>interpretability</i>)
	Verständlichkeit (<i>understandability</i>)
	Konsistenz der Darstellung (<i>representational consistency</i>)
	Knappheit der Darstellung (<i>representational conciseness</i>)
Zugriffsqualität	Verfügbarkeit (<i>accessibility</i>)
	Zugriffssicherheit (<i>access security</i>)

Quelle: Naumann (2007), S. 29.

Die Bundesnetzagentur hat für die Berechnungen des Törnqvist-Index im Wesentlichen zwei Datenquellen verwendet: Zum einen die Daten der Unternehmen, die aus den jeweiligen Abfragen hervorgingen, zum anderen Indexreihen des Statistischen Bundesamtes.

In Bezug auf die Glaubhaftigkeit können beide Datenquellen als einwandfrei bezeichnet werden. Beim Aspekt der Genauigkeit ist zu beachten, dass es sich bei den Unternehmensdaten um eine nahezu vollständige Erhebung handelt, während die Daten des Statistischen Bundesamtes überwiegend auf Stichproben beruhen. Es kann davon ausgegangen werden, dass die Objektivität bei beiden Datenquellen in gleicher Weise gegeben ist, da sie letztendlich beide aus den Antworten der Unternehmen resultieren. Dasselbe gilt für die Reputation.

Die übrigen Merkmale (kontextuelle und repräsentative Datenqualität sowie Zugriffsqualität) spielen bei der Auswahl einzelner konkreter Datensätze eine Rolle und werden somit im nächsten Abschnitt 4.2 bei der jeweiligen Methodik wieder aufgegriffen, soweit dies hilfreich erscheint.

4.2 Methodik

In diesem Abschnitt wird auf Aspekte eingegangen, die, wie sich in der laufenden Regulierungsperiode gezeigt hat, auf die Berechnung des generellen sektoralen Produktivitätsfortschritts einen wesentlichen Einfluss besitzen.

4.2.1 Stützintervall

Die Ermittlung des Xgen hat gemäß § 9 Abs. 3 S. 2 ARegV unter Einbeziehung der Daten von Netzbetreibern aus dem gesamten Bundesgebiet für einen Zeitraum von mindestens vier Jahren zu erfolgen. Für die Festlegung des Xgen für die dritte Regulierungsperiode hat die Bundesnetzagentur für den Gasbereich den Zeitraum von 2006 bis 2016 und für den Strombereich von 2006 bis 2017 gewählt.

Eine allgemeine Diskussion zum optimalen Stützintervall findet sich bereits in Abschnitt 2.3.2. An dieser Stelle wird daher auf zwei konkrete Aspekte eingegangen. Zum einen wird der Umgang mit Stützintervallen im internationalen Kontext an zwei Beispielen aufgezeigt. Zum anderen wird der sog. Basisjahreffekt diskutiert.

In den **Niederlanden** wurde für die Verteilnetzbetreiber für die Regulierungsperiode 2017 bis 2021 ein ähnlicher Zeitraum angesetzt wie in Deutschland. Für den Gassektor wurde das Stützintervall von 2006 bis 2015 gewählt.¹⁰⁵ Für die Regulierungsperiode 2022 bis 2026 wurde der Zeitraum bis 2020 verlängert, das Stützintervall spannt sich nun also zwischen den Jahren 2006 und 2020 auf.¹⁰⁶ Für den Stromsektor wurde für die aktuelle Regulierungsperiode (2022 bis 2026) das Stützintervall 2005 bis 2020 gewählt.¹⁰⁷

¹⁰⁵ ACM (2021b), Rn. 306.

¹⁰⁶ ACM (2021b), Rn. 304.

¹⁰⁷ ACM (2021a), Rn. 267.

Die Begründung für einen möglichst langen Betrachtungszeitraum ist für die Sektoren Strom und Gas ähnlich. So heißt es im Methodenbeschluss für die Stromverteilernetzbetreiber:¹⁰⁸

*„Nach Ansicht der ACM erfüllt die langfristige Produktivitätsveränderung am besten die Kriterien, die für einen guten Schätzer der Produktivitätsveränderung gelten, nämlich Repräsentativität und Robustheit. Auch die ACM ist der Ansicht, dass eine erneute Nutzung der langfristigen Produktivitätsveränderung einer effektiven Anreizwirkung für das Regu-
lierungsverfahren insgesamt nicht entgegensteht. Schließlich sieht ACM in der Verfügbarkeit von mehr Daten der Netzbetreiber einen weiteren Grund, sich für die langfristige Produktivitätsänderung zu entscheiden.“*

Für den Gassektor stellt ACM fest:¹⁰⁹

„Je repräsentativer die verwendeten vergangenen Erkenntnisse und je robuster der Schätzer ist, desto genauer wird die Schätzung der zukünftigen Produktivitätsänderung sein, die regionale Netzbetreiber erfahren könnten. Repräsentativ bedeutet hier, dass erwartet wird, dass die verwendeten vergangenen Realisierungen mit zukünftigen Realisierungen vergleichbar sind, beispielsweise weil die Umstände im Bewertungszeitraum mit denen in dem Zeitraum vergleichbar sind, für den die Änderung der Produktivität geschätzt wird. Robust bedeutet, dass die endgültige Schätzung der Produktivitätsänderung auf möglichst vielen repräsentativen Erkenntnissen aus der Vergangenheit basiert. Dadurch werden die Auswirkungen von Zwischenfällen oder Messfehlern auf die Produktivitätsänderung minimiert.“

In **Großbritannien** wurde für Strom und Gas der Zeitraum von 1997 bis 2016 gewählt.¹¹⁰ Im Methodenpapier für Ofgem wird dazu ausgeführt:¹¹¹

„Wir empfehlen, dass der geeignete Zeitraum für die Berücksichtigung der Growth Accounting-Analyse 1997 bis 2016 ist. Unsere Begründung für diese Empfehlung ist, dass der vorgeschlagene Zeitraum zwei vollständige Geschäftszyklen umfasst, was das Maximum ist, das im EU-KLEMS-Datensatz 2019 enthalten ist. Wir sind der Ansicht, dass ein längerer Zeitraum über mindestens zwei Geschäftszyklen wahrscheinlich die Empfindlichkeit gegenüber Messfehlern verringert. Dies wird zu einer durchschnittlichen Produktivitätswachstumsrate führen, die möglicherweise repräsentativer für langfristige zugrunde liegende Faktoren ist.“

Beide Länder betrachten somit einen möglichst langen Betrachtungszeitraum als wesentlichen Aspekt bei der Wahl des Stützintervalls. Weiterhin sollte dieser die wahrscheinliche Entwicklungsumgebung der zu regulierenden zukünftigen Periode möglichst gut wiedergeben. Die Bundesnetzagentur ist in ihren Festlegungen einer solchen Rationalität ebenfalls gefolgt.

Im Zuge der Festlegungen wurde auch über den sog. **Basisjahreffekt** diskutiert. Zu Beginn einer Regulierungsperiode werden die Kosten der Netzbetreiber auf das aktuelle

¹⁰⁸ ACM (2021a), Rn. 264, eigene Übersetzung.

¹⁰⁹ ACM (2021b), Rn. 295, eigene Übersetzung.

¹¹⁰ Ofgem (2020), S. 49, CEPA (2020a), S. 6.

¹¹¹ CEPA (2020a), S.17.

netzbetreiberindividuelle Niveau eingerastet. Der Basisjahreffekt tritt somit ein, wenn die Netzbetreiber Investitionen in das Jahr verschieben, das für die Bestimmung der Erlösobergrenze maßgeblich ist (das Basisjahr). Es können sowohl Investitionen in das Basisjahr nach vorne verschoben werden (aus dem Jahr oder den Jahren vor dem Basisjahr), als auch vorgezogen werden (aus dem Jahr oder den Jahren nach dem Basisjahr). Wenn der Output in diesen Jahren vergleichsweise gleich hoch bleibt, ergibt sich insbesondere für die beiden Jahre vor und nach dem Basisjahr ein geringerer Input in Form geringerer Kosten, die Produktivität steigt an.

Vor diesem Hintergrund wurde diskutiert, ob das Jahr 2006 (als Basisjahr) noch in das Stützintervall aufzunehmen sei, da bei seiner Aufnahme die beschriebenen Effekte rund um das Basisjahr nicht vollständig abgebildet würden. Andererseits ist es einer längerfristigen Betrachtung immanent, dass in diesen Zeitraum fallende Ereignisse mit eingehen. Gerade hierdurch wird die Unterschiedlichkeit möglicher Entwicklungen mitberücksichtigt und die Prognosegüte gesteigert. Dabei werden, je länger der Betrachtungszeitraum ist, die unterschiedlichen Entwicklungen entsprechend geglättet. Wie oben beschrieben wird auch im internationalen Kontext vor dem Hintergrund der Abbildung der langen Investitionszyklen generell angestrebt, einen möglichst langen Zeitraum als Stützintervall zu wählen, solange die betrachteten Jahre eine Prognose für die Zukunft erlauben.

Der generelle sektorale Produktivitätsfaktor dient dazu, im Rahmen der Erlösobergrenzenformel den VPI hinsichtlich etwaiger Besonderheiten der Netzbranche zur Einstands- und Produktivitätsentwicklung zu korrigieren. Dabei soll die Anreizregulierung grundsätzlich Wettbewerb simulieren.

Mithin ist es das Kalkül der Netzbetreiber, Investitionen in das Basisjahr zu verschieben, um die erlaubten Erlöse zu erhöhen. Wirkt sich dieses Verhalten im Rahmen der Xgen-Festlegung nun zu Ungunsten der Netzbetreiber aus, so steht dies der regulierungsbehördlichen Entscheidung für das längstmögliche Stützintervall nicht entgegen. Vielmehr sind auch Unternehmen im Wettbewerb verschiedenen Risiken unterworfen, mit denen sie umgehen müssen.

4.2.2 Outputfaktoren

Als Outputfaktoren zur Berechnung der Produktivität kommen verschiedene Parameter in Frage. Die Bundesnetzagentur hat in ihrer Entscheidung für die dritte Regulierungsperiode die deflationierten Umsatzerlöse als Output herangezogen.

Vor der entsprechenden Festlegung wurde dies durch ein entsprechendes Gutachten für den Strombereich untersucht. Darin heißt es unter anderem:¹¹² „Gestützt auf unsere Analysen und in Beachtung der verfügbaren Daten und deren Qualität empfehlen wir, auch im Strom den Bruttoproduktionswert zu verwenden, um den Xgen der RP3 festzulegen.“

¹¹² Trinkner und Mattmann (2018), S. 39.

Der Bundesgerichtshof hat in seinem Beschluss zum generellen sektoralen Produktivitätsfaktor im Gasbereich festgehalten:¹¹³ „Von Rechts wegen nicht zu beanstanden ist es ferner, dass und in welcher Weise die Bundesnetzagentur den Bruttoproduktionswert des Netzbetriebs als Outputfaktor im Rahmen des von ihr verwendeten Törnqvist-Index zugrunde gelegt hat.“

Im Folgenden werden dennoch verschiedene Optionen für mögliche Outputparameter nochmals aufgegriffen und dargelegt, sowie deren Vor- und Nachteile aufgeführt.

Physische Outputfaktoren

Da es sich um eine Produktivitätsbetrachtung handelt, können grundsätzlich zunächst auch nichtmonetäre Outputfaktoren für die Berechnung herangezogen werden. Dies können zum Beispiel der Gas- oder Stromabsatz sein oder auch die Anzahl der angeschlossenen Kunden. Allerdings sollte es sich um einen oder mehrere möglichst netznahe Parameter handeln, um den Sachverhalt bestmöglich abzudecken.

Der physische Output der Netzbetreiber ist normalerweise nicht eindimensional. Outputs können verschiedene Größen sein, so wie etwa auch für den Effizienzvergleich (vgl. § 13 Abs. 3 ARegV):

- die Anzahl der Anschlusspunkte oder der Zählpunkte in Stromversorgungsnetzen und der Ausspeisepunkte oder der Messstellen in Gasversorgungsnetzen,
- die Fläche des versorgten Gebietes,
- die Leitungslänge oder das Rohrvolumen,
- die Jahresarbeit,
- die zeitgleiche Jahreshöchstlast,
- die dezentralen Erzeugungsanlagen in Stromversorgungsnetzen.

Falls mehrere dieser Größen als Outputvariablen herangezogen werden, müssen sie in geeigneter Form aggregiert, d. h. entsprechend gewichtet werden. Dies kann auf Basis ingenieurwissenschaftlicher Zusammenhänge oder der Wertanteile der einzelnen Parameter erfolgen.

Um die Wertanteile zu bestimmen, müssten allerdings für jeden Output die Umsatzerlöse vorliegen, um so die Wertquoten (Teilumsatz des Outputs dividiert durch gesamte Umsatzerlöse) zu bilden. Während dies für die Outputfaktoren Last und Arbeit darstellbar erscheint, sind diese für andere Outputs, die herangezogen werden müssten, um den Output der Netzbetreiber vollständig abzubilden, nicht vorhanden, da keine Preise für diese erhoben werden.

Eine Prüfung bzw. Validierung der Gewichtung eines Ansatzes im Bereich Strom mit den Parametern Last, Arbeit, Ein- und Ausspeisepunkte sowie Stromkreislängen auf Basis

¹¹³ Bundesgerichtshof (2021), Rn. 45.

von Paneldaten hat gezeigt, dass kein robustes Modell mit durchweg positiven Koeffizienten ermittelbar ist.¹¹⁴

Das Heranziehen physischer Faktoren ist somit auch mit Schwierigkeiten und Unsicherheiten behaftet.

Bruttowertschöpfung vs. Produktionswert

Ein Vorteil monetärer Größen ist es, dass die Werte nicht einzelnen Outputs zugeordnet werden müssen. Vielmehr bilden sie sämtlichen Output der Netzbetreiber ab.

Neben dem Produktionswert (den die Bundesnetzagentur in der dritten Regulierungsperiode sowohl für den Strombereich als auch für den Gasbereich in Form der deflationierten Umsatzerlöse herangezogen hat) steht auch die Bruttowertschöpfung als Outputfaktor zur Verfügung.¹¹⁵ Der Produktionswert zieht neben den Produktionsfaktoren Arbeit und Kapital auch die Vorleistungen mit ein. Die Gewichtungen der einzelnen Produktionsfaktoren entsprechen dabei ihrem Anteil am Produktionswert. Die Summe der Gewichtungen addiert sich zu eins. Bei der Bruttowertschöpfung hingegen werden die Vorleistungen nicht berücksichtigt, sie entspricht somit dem Produktionswert abzüglich eben dieser Vorleistungen.

Vorleistungen können teilweise nicht eindeutig vom Kapitaleinsatz abgegrenzt werden, was bei längeren Zeitreihen aufgrund von im Zeitablauf unterschiedlichen Definitionen zu Problemen führen kann. Entsprechende Änderungen der Vorleistungsquote bedingen dann eine Veränderung der Produktivität, ohne dass diese real erfolgt ist.

Diese Problematik stellt sich bei Anwendung der Bruttowertschöpfung nicht dar. Allerdings ist die Bruttowertschöpfung eine abgeleitete Größe, so dass sie größeren Schwankungen unterliegen kann. Weiterhin wird das Outsourcing von Leistungen nicht mit in die Berechnung einbezogen. Somit empfiehlt es sich zunächst grundsätzlich beide Größen für eine Berechnung zu verwenden.

Bei der Berechnung der TFP für eine bestimmte Industrie ergibt sich auf Grundlage der Bruttowertschöpfung allerdings üblicherweise, wenn technologischer Fortschritt alle Inputfaktoren in einem proportionalen Verhältnis betrifft, ein systematisch höherer Wert als auf Grundlage des Produktionswertes.¹¹⁶

Aus diesem Grund kann es aus Vorsichtsgründen geboten sein, den Produktionswert der Bruttowertschöpfung vorzuziehen. Unter der Voraussetzung also, dass technischer Fortschritt alle Inputfaktoren proportional betrifft, stellt die Produktivitätsmessung auf Basis des Produktionswertes die aus regulatorischer Sicht vorteilhaftere Methode dar.¹¹⁷ Dies wird auch in der Wissenschaft so beschrieben.

¹¹⁴ Trinkner und Mattmann (2018), S. 35ff.

¹¹⁵ Vgl. im Folgenden Liebe et al. (2017), S. 42.

¹¹⁶ Vgl. dazu weitergehende Ausführungen in: OECD (2001), S. 27 ff.

¹¹⁷ Vgl. dazu beispielsweise OECD (2001): S. 27 oder Oulton (2017): „The mystery of TFP“, S. 18.

Der **Produktionswert** wird in zahlreichen Studien zur Produktivitätsberechnung als Output herangezogen. Im Folgenden sind dazu einige Beispiele aufgeführt:

- Bils, M., Klenow, P. and Ruane, C. (2020): Misallocation or Mismeasurement? NBER Working Paper No. w26711.
In diesem Papier wird der Produktionswert (Gross Output) als Basis von Produktivitätsberechnungen auf Unternehmensebene für das produzierende Gewerbe in Indien und den USA herangezogen.
- Nazarkoa, J. und E. Chodakowska (2015): Measuring productivity of construction industry in Europe with Data Envelopment Analysis, in: Procedia Engineering 122 (2015), S. 204 – 212.
In dieser Untersuchung wird u. A. der Umsatz als Outputgröße herangezogen.
- Oberfield, E. (2013) "Productivity and Misallocation During a Crisis: Evidence from the Chilean Crisis of 1982," Review of Economic Dynamics, 2013, 16 (1), S.100–119.
In diesem Artikel wurden Unternehmensdaten aus der chilenischen Erhebung bezüglich des produzierenden Gewerbes verwendet, um den Rückgang der gemessenen TFP zu untersuchen. Dazu wurde der Brutto-Output zugrunde gelegt.
- Huang, K.S. (2002): Measuring Food Manufacturing Productivity: Gross- or Net-Output Approach?, prepared for presentation at the American Agricultural Economics Association 2002 Annual Meeting in Long Beach, CA, July 28-31.
„Der aus dem Brutto-Output-Ansatz ermittelte TFP-Index wird als repräsentatives Produktivitätsmaß für den US-amerikanischen Lebensmittelsektor angesehen.“(eigene Übersetzung aus dem Englischen)

Das Statistische Bundesamt ermittelt in der VGR sowohl den Produktionswert als auch die Bruttowertschöpfung. Zum Zwecke der Produktivitätsmessung eines Sektors erscheint der Produktionswert deswegen vorteilhafter, da er alle Inputs einschließt. Auf Sektorebene kann davon ausgegangen werden, dass nur wenige Produkte und Dienstleistungen des eigenen Sektors in die Produktion eingehen. Daher bildet der Produktionswert ein adäquates Maß.

Auf gesamtwirtschaftlicher Ebene ergibt sich dagegen ein anderes Bild. Hier fließen Güter und Dienstleistungen verschiedener Sektoren als Vorleistungen in andere Sektoren ein (ausgenommen Importe). Misst man also die Produktivität der Gesamtwirtschaft, so bietet sich hier eher die Bruttowertschöpfung als Maßstab an, um Doppelzählungen zu vermeiden.¹¹⁸

Internationale Erfahrungen

Ein Blick in andere Länder zeigt, dass die Berechnung der Produktivität auf Basis von Wertgrößen durchaus üblich ist. Der britische Regulierer Ofgem beispielsweise basiert die Berechnungen des Frontier-Shifts auf EU-KLEMS-Daten.¹¹⁹ Dazu werden die Bruttowertschöpfung bzw. der Produktionswert als Outputparameter herangezogen.¹²⁰ Diese Vorgehensweise fand seitens der Branche keinerlei Einwände, vielmehr entfaltete sich die Diskussion darum, welche der beiden wertmäßigen Größen heranzuziehen sei.¹²¹

¹¹⁸ Vgl. hierzu z. B. Blum et al. (1999), S. 64.

¹¹⁹ Vgl. euklems.eu, zuletzt abgerufen am 06.04.2022.

¹²⁰ CEPA (2020b), S.11 ff.

¹²¹ CEPA (2020b), NERA (2020).

Die niederländische Regierungsbehörde ACM verwendet einen zusammengesetzten Output als sog. Leistung der Netzbetreiber. Dies bedarf einer eindeutigen Bewertung und Zuordnung *der Umsätze* jedes Netzbetreibers zu verschiedenen Outputparametern. Beispiele im Bereich der Gasverteilnetzbetreiber sind die Kapazität der Anschlüsse (in Kubikmeter Gas pro Stunde und Jahr) und die Anzahl der Kunden.¹²² Ein Nachteil ist dabei, dass sich durch Korrekturen der Erlöse die Gewichte der Parameter ändern können.¹²³

Vor- und Nachteile verschiedener Parameter

Die Vor- und Nachteile der in Frage kommenden Outputfaktoren werden in Tabelle 4-1 noch einmal aufgeführt.

Tabelle 4-1: Vor- und Nachteile verschiedener Outputfaktoren

Outputfaktor	Vorteile	Nachteile
Physische Outputfaktoren	<ul style="list-style-type: none"> • Grundsätzlich einfach zu erheben 	<ul style="list-style-type: none"> • Gewichtung nötig, falls mit anderen Faktoren kombiniert • Umsatzerlöse pro Parameter notwendig • Abgrenzung der Parameter schwierig • Anfällig für Datenfehler / Korrekturbedarf
Bruttowertschöpfung	<ul style="list-style-type: none"> • Relativ einfache Erfassung verschiedener Outputs • Zuordnung der Erlöse zu einzelnen Outputs entfällt 	<ul style="list-style-type: none"> • Monetäre Größe, muss deflationiert werden • Erfasst keine Vorleistungen • Größeren Schwankungen unterworfen
Produktionswert	<ul style="list-style-type: none"> • Relativ einfache Erfassung verschiedener Outputs • Zuordnung der Erlöse zu einzelnen Outputs entfällt • Bildet sektorinterne Umstrukturierungen besser ab • Konservativer Ansatz 	<ul style="list-style-type: none"> • Monetäre Größe, muss deflationiert werden

Quelle: Eigene Darstellung

Im Ergebnis ist festzuhalten, dass alle Verfahren bereits Anwendung gefunden haben und sich eine klare Vorteilhaftigkeit eines potenziellen Outputfaktors mithin nicht ergibt. Aus Vorsichtsgründen kann aber bei der Bestimmung des Xgen auf den Produktionswert zurückgegriffen werden.

4.2.3 Deflatoren

Produktivitätsberechnungen basieren auf realen Werten. Daher müssen nominelle monetäre Werte, wenn sie zur Berechnung herangezogen werden, deflationiert werden, so dass potenzielle Preisänderungen nicht als Produktivitätsänderungen interpretiert werden können. Aufgabe eines Deflators ist es also, die Umwandlung von nominalen in reale Größen zu ermöglichen. Ein hoher (niedriger) Deflator führt dabei zu einem niedrigeren

¹²² ACM (2021b), S. 33.

¹²³ ACM (2015), Rn 149.

(höheren) realen Wert. Das bekannteste Beispiel eines Deflators ist der sogenannte BIP-Deflator, mithin definiert als das Verhältnis des nominalen zum realen BIP.¹²⁴

In der dritten Regulierungsperiode wurden für die Bildung des Produktionswerts die Umsatzerlöse, die Bestandsveränderungen sowie die aktivierten Eigenleistungen herangezogen. Diese wurden jeweils mit eigenen Deflatoren preisbereinigt. Für die Umsatzerlöse, die sowohl im Strom- als auch im Gasbereich mehr als 98 % des Produktionswerts ausmachen, wurde ein Deflator auf Grundlage der Netzentgelte für verschiedene Kundengruppen (Haushalte, Gewerbe, Industrie) gebildet, wobei die jeweiligen Gewichtungen dieser Gruppen über den Betrachtungszeitraum fixiert waren. Dabei wurde ein Durchschnittswert über alle Jahre gebildet, so dass einerseits der aktuellen Entwicklung Rechnung getragen wurde, andererseits aber auch eine Anlehnung an das Vorgehen von Statistikbehörden stattfand.

Der Deflator soll dabei die durchschnittliche Preisentwicklung der Umsatzanteile abbilden. Die Deflationierung der aktivierten Eigenleistungen sowie der Bestandsveränderungen erfolgte auf Basis von Zeitreihen des Statistischen Bundesamtes. Dies gilt ebenso für die verschiedenen verwendeten Aufwendungsgrößen.

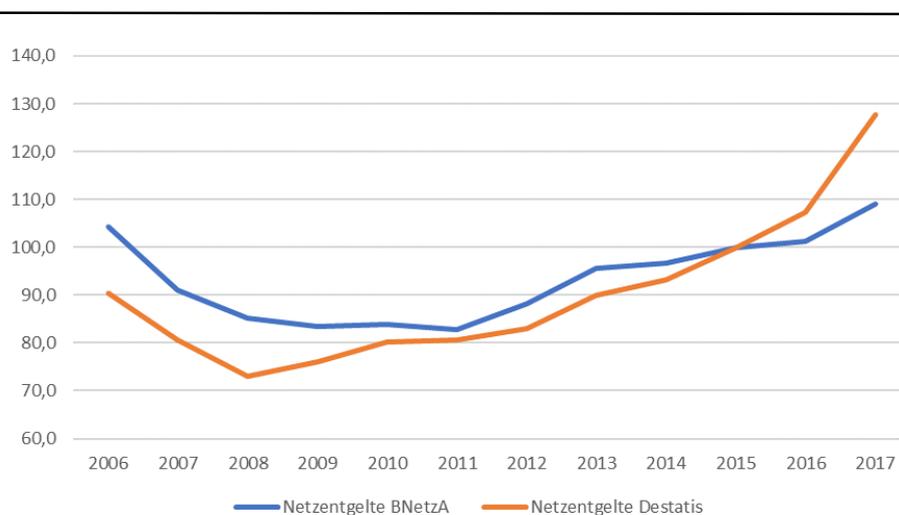
Als Grundlage für die Netzentgelte zur Deflationierung der Umsatzerlöse hat die Bundesnetzagentur die Netzentgelte aus den Monitoringberichten entnommen. Diese Daten werden von den Lieferanten erhoben, sodass auf Grundlage dieser Vollerhebung die Entwicklung der Netzentgelte bei allen Netzbetreibern in Deutschland berücksichtigt werden. Grundsätzlich soll der Deflator an dieser Stelle die tatsächlichen Umsätze sämtlicher Netzbetreiber vorgelagerter Netzebenen erfassen. Dabei kann durchaus berücksichtigt werden, dass die Netzentgelte höherer Spannungsebenen an untere Netzebenen weitergegeben werden (sogenannte Kaskadierung).

Als Alternative könnten hier auch entsprechende Deflatoren des Statistischen Bundesamtes (Destatis) herangezogen werden. Die Netzentgelte für Strom umfassen die Dienstleistung der Elektrizitätsübertragung, der Elektrizitätsverteilung und des Elektrizitätshandels.¹²⁵ Ein Vergleich dieser Zeitreihe mit dem von der Bundesnetzagentur angewandten Netzentgeltdeflator zeigt allerdings, dass diese sich in ihrer Entwicklung nicht wesentlich unterscheiden (vgl. Abbildung 4-2), aber im Detail wichtige systematische Differenzen aufweisen. So gehen in die Netzentgeltbildungen der nachgelagerten Ebenen bspw. die vorgelagerten Netzkosten mit ein (sog. Kaskadierungseffekte), was schier unauflösbare Gewichtungprobleme hervorruft. Ferner entfalten die Werte der oberen Spannungsebenen keine tatsächliche Anwendungswirkung, da die direkt angeschlossenen Netzkunden über Speicherbefreiungs- oder auch Poolingregelungen oder § 19 StromNEV-Rabattierungen regelmäßig nur Bruchteile der ausgewiesenen Preise bezahlen. Ab 2012 wurde durch die § 19-StromNEV-Umlage zudem das Netzkostenvolumen entlastet. Eine direkte Einsatzmöglichkeit der Werte des statistischen Bundesamtes erscheint daher fraglich.

¹²⁴ Vgl. z. B. Mankiw und Taylor (2018), S.804.

¹²⁵ Für Gas werden die Netzentgelte dort nicht direkt ausgewiesen.

Abbildung 4-2: Netzentgeltdeflatoren Strom (Index 2015 = 100)



Quelle: Eigene Darstellung auf Basis von Bundesnetzagentur (2018) und Statistisches Bundesamt (2022), S. 218.

Gleichwohl wird sichtbar, dass die Entwicklung nahezu parallel verläuft, während die Netzentgeltreihe der Bundesnetzagentur bis zum Jahr 2015 über dem Index von Destatis liegt, der nominale Output für die betreffenden Jahre also stärker deflationiert wird, was (im Vergleich) einen geringeren Xgen zur Folge hätte. Die Korrelationsanalyse der beiden Zeitreihen zeigt einen starken Zusammenhang (Pearson-Korrelationskoeffizient vom 0,875), der statistisch signifikant ist.

Weiterhin hätte auch die Anwendung des Destatis-Deflatoren verschiedene Schwächen: zum einen beruhen die Erhebungen auf Stichproben bei den Netzbetreibern, während der Bundesnetzagentur-Ansatz auf Basis einer Vollerhebung erfolgt. Zudem wird bei Destatis auf die Standardpreisblätter rekurriert, während der Deflator auf Basis des Monitoring-Index auf den tatsächlich gezahlten Netzentgelten beruht. Weiterhin ist festzuhalten, dass die Sondernetzentgelte nach § 19 Abs. 2 S. 2 StromNEV zu einer tatsächlichen Reduktion der tatsächlich gezahlten Netzentgelte und damit zu einer von den deutlich steigenden Netzentgelten des Standardpreissystems abweichenden Entwicklung geführt haben. Auch wurden einige Verbraucher aus dem System der Netzentgelte herausgenommen. So zahlten Pumpspeicherkraftwerke und andere Speicher in den ersten Jahren des hier relevanten Zeitraumes noch Netzentgelte, wurden aber später gänzlich von den Netzentgelten befreit.

An dieser Stelle bleibt also ein Entscheidungsspielraum für die Regulierungsbehörde und es spricht einiges dafür, den genaueren Ansatz der Bundesnetzagentur zu favorisieren. Jedenfalls ist eine eindeutige Überlegenheit eines bestimmten Ansatzes nicht erkennbar.

Tabelle 4-2: Vor- und Nachteile verschiedener Deflatoren

Deflator	Vorteile	Nachteile
Netzentgelte auf Basis der Monitoring-Berichte	<ul style="list-style-type: none"> • Daten einfach verfügbar • Vollerhebung • Tatsächliche Netzentgelte 	<ul style="list-style-type: none"> • Höhere Spannungsebenen nicht direkt erfasst
Destatis-Index	<ul style="list-style-type: none"> • Daten einfach verfügbar 	<ul style="list-style-type: none"> • Stichprobe • Daten beruhen auf Standardpreisblättern ((Anwendungsveränderungen und Rabatte nicht berücksichtigt) • Gewichtung erhält Erlöse aus Weiterverteilung an nachgelagerte Netzbetreiber

Quelle: Eigene Darstellung

4.2.4 Inputpreisreihen

Da ein Inputpreisindex weder für das Gas- noch das Stromnetz verfügbar ist, hat die Bundesnetzagentur einzelne Inputpreisindizes herangezogen und entsprechend der Anteile in der GuV gewichtet. Die Auswahl der Indizes wird im Folgenden diskutiert und mögliche Alternativen aufgezeigt.

Die *Personalkosten* wurden sowohl für Strom als auch für Gas aus den abgefragten Daten der Netzbetreiber (GuV) als Quotient aus Personalaufwand und tatsächlich geleisteten Arbeitsstunden berechnet. Als Alternative könnte der Arbeitskostenindex der Energieversorgung von Destatis verwendet werden.

Die Abfrage der Personalkosten durch Destatis beruht allerdings nur auf Stichproben bei den Unternehmen¹²⁶, während die Bundesnetzagentur eine Vollerhebung durchgeführt hat. Dabei ergibt sich die Stärke des von der Bundesnetzagentur gewählten Ansatzes aus der besseren Datengrundlage durch die Vollerhebung und der Tatsache, dass es sich ausschließlich um Netzbetreiberdaten handelt, während Destatis die gesamte Energieversorgung einschließt.

Die Preisentwicklung für die *Aufwendungen für Roh-, Hilfs- und Betriebsstoffe* werden durch den „Index der Erzeugerpreise gewerblicher Produkte (Vorleistungsgüterproduzenten und Investitionsgüterproduzenten)“ und den „Index der Erzeugerpreise Erdgas, bei Abgabe an Wiederverkäufer“ (für den Gasbereich) bzw. „Elektrischer Strom an Weiterverteilern“ (für den Strombereich) des Statistischen Bundesamtes abgebildet.

Die Indizes für Vorleistungsgüter und Investitionsgüter sind Teil einer separaten Aufgliederung des Index der Erzeugerpreise gewerblicher Produkte. Neben den beiden genannten Gütern enthält dieser auch Konsumgüterproduzenten, Gebrauchsgüterproduzenten, Verbrauchsgüterproduzenten und Energie.¹²⁷ Die Indizes enthalten u. a. Metalle, Metall-erzeugnisse, elektrische Ausrüstungen und Maschinen. Das Heranziehen noch tieferer

¹²⁶ Statistisches Bundesamt (2021), S. 4.

¹²⁷ Statistisches Bundesamt (2019), Kapitel 4, S. 16.

Gliederungsebenen des Erzeugerpreisindex und somit die Auswahl von Einzelindizes wäre theoretisch möglich. Dies würde aber die Granularität der Datenabfrage weiter erhöhen, um die Indizes entsprechenden Positionen der Unternehmen zuordnen zu können. Ob damit die Qualität der Berechnung verbessert oder nur eine Scheingenauigkeit erzielt würde, bleibt zumindest zweifelhaft.

Für die Kostenposition „Betriebsverbrauch“ werden die beiden oben genannten Indizes herangezogen. Für den Strombereich handelt es sich um den Preis ab Kraftwerk. Der Betriebsverbrauch ist der Betrag, den der Netzbetreiber zur eigenbetrieblichen Nutzung verwendet.¹²⁸ Für den Gasbereich erscheint der genannte Index geeignet, da „er den Preis wiedergibt, der durch die Erdgasförderer bzw. Importeure erhoben wird. Die Zwischenstufe der Wiederverkäufer (und deren Aufschlag) entfällt somit.“¹²⁹

Die Position *Aufwendungen für bezogene Leistungen* (exklusive der Position vorgelagerte Netzkosten) setzt sich aus Aufwendungen für durch Dritte erbrachte Betriebsführung sowie Wartungs- und Instandhaltungsaufwendungen und auch aus Aufwendungen für die Beschaffung von Differenzmengen und Ausgleichenergie zusammen.¹³⁰ Als Indizes wurden die Arbeitskostenindizes Produzierendes Gewerbe und Dienstleistungsbe-
reich (WZ08-B-08) sowie der Index der Erzeugerpreise Erdgas bei Abgabe an Wiederverkäufer¹³¹ bzw. Elektrischer Strom an Weiterverteiler¹³² herangezogen. Die Festlegung der jeweiligen Anteile erfolgte auf Basis der Daten der jeweiligen Kostenprüfungen.

Der Arbeitskostenindex WZ08-B-08 umfasst nach der Klassifikation des Statistischen Bundesamtes alle Wirtschaftszweige von B bis S, wie in Tabelle 4-3 dargestellt.

¹²⁸ Vgl. Liebe et al. (2017), S.88.

¹²⁹ Liebe et al. (2017), S.94.

¹³⁰ Bundesnetzagentur (2018a), S. 23, Bundesnetzagentur (2018b), S. 31.

¹³¹ Bundesnetzagentur (2018a), S. 24.

¹³² Bundesnetzagentur (2018b), S. 31.

Tabelle 4-3: Wirtschaftszweige im Arbeitskostenindex Produzierendes Gewerbe und Dienstleistungsbereich

WZ	Wirtschaftszweig
B	Bergbau und Gewinnung von Steinen und Erden
C	Verarbeitendes Gewerbe
D	Energieversorgung
E	Wasserversorg.; Entsorg., Beseit. v. Umweltverschmutz.
F	Baugewerbe
G	Handel; Instandhalt. u. Reparatur von Kraftfahrzeugen
H	Verkehr und Lagerei
I	Gastgewerbe
J	Information und Kommunikation
K	Erbringung von Finanz- u. Versicherungsdienstleistungen
L	Grundstücks- und Wohnungswesen
M	Erbring. v. freiberufl., wiss. u. techn. Dienstleistung
N	Erbringung v. sonst. wirtschaftl. Dienstleistungen
O	Öffentl. Verwaltung, Verteidigung; Sozialversicherung
P	Erziehung und Unterricht
Q	Gesundheits- und Sozialwesen
R	Kunst, Unterhaltung und Erholung
S	Erbringung von sonstigen Dienstleistungen

Quelle: Eigene Darstellung auf Basis Statistisches Bundesamt (2017), S. 43 ff.

Wie zu sehen ist, bezieht der Index auch Leistungen der Energieversorgung mit ein. Als Alternative könnte dieser Wirtschaftszweig aus dem Arbeitskostenindex herausgenommen werden, da er weitestgehend über die Preise für Strom und Gas abgedeckt wird. Allerdings spricht die geringe Gewichtung der beiden letztgenannten Indizes (1 % im Strombereich und 7 % im Gasbereich) bei der Zusammensetzung eines Index für „Aufwendungen für bezogene Leistungen“ dafür, dass eine solche Vorgehensweise nur sehr geringe Auswirkungen hätte und daher wissenschaftlich nicht eindeutig vorzuziehen wäre.

Die Preise für *sonstige betriebliche Aufwendungen* werden jeweils durch einen Mischindex abgebildet. Es wurde versucht, den verschiedenen Positionen der Netzbetreiber durch das Heranziehen entsprechender Zeitreihen des Statistischen Bundesamtes bestmöglich gerecht zu werden. Da die sonstigen betrieblichen Aufwendungen sehr unterschiedlich ausfallen, wurden verschiedene Zeitreihen genutzt. Da entsprechende Daten bei den Netzbetreibern nicht abgefragt wurden, kann diese Herangehensweise nachvollzogen werden. Einer Systematisierung hätten sich diese Einzelabfragen darüber hinaus vermutlich entzogen, da die Bandbreite möglicher Positionen im dreistelligen Bereich liegen dürfte.

Im Laufe der Verfahren wurde kritisiert, dass nicht für alle Kostenpositionen des Erhebungsbogens eine entsprechende Berücksichtigung bei der Xgen-Berechnung gefunden

hätten. So sei z. B. „Wartung und Instandsetzung“ nicht ausreichend berücksichtigt worden. Hierbei wird allerdings verkannt, dass der herangezogene Arbeitskostenindex Produzierendes Gewerbe und Dienstleistungsbereich (WZ08-B-08) bereits entsprechende Positionen enthält. Dies geschieht unter anderem in den folgenden Kategorien:¹³³

- „Gruppe 33.1 ("Reparatur von Metallerzeugnissen, Maschinen und Ausrüstungen")
- Abteilung 43 ("Vorbereitende Baustellenarbeiten, Bauinstallation und sonstiges Ausbaugewerbe")
- Gruppe 45.2 ("Instandhaltung und Reparatur von Kraftwagen")
- Abteilung 95 ("Reparatur von Datenverarbeitungsgeräten und Gebrauchsgütern")

Somit kann der Ansatz aus wissenschaftlicher Sicht nicht beanstandet werden, zumal nicht ersichtlich ist, welche Indexreihe an dieser Stelle alternativ hätte verwendet werden können.

Für *Zinsen und ähnliche Aufwendungen* hat sich die Bundesnetzagentur entschieden, analog § 7 Abs. 7 der Gas- bzw. Stromnetzentgeltverordnung vorzugehen. Dieser Zinssatz setzt sich zusammen aus den Umlaufrenditen inländischer Inhaberschuldverschreibungen/Anleihen der öffentlichen Hand (Monatsdurchschnitte), der Umlaufrenditen inländischer Inhaberschuldverschreibungen / Hypothekendarlehen (Monatsdurchschnitte) sowie den Umlaufrenditen inländischer Inhaberschuldverschreibungen / Anleihen von Unternehmen (Nicht-MFIs).

Kritik zu diesem Posten bezog sich auf die Art der Mittelwertbildung (siehe dazu den Abschnitt zu Fremdkapitalzinsen in Abschnitt 4.2.6), nicht auf die Auswahl der Indizes. Die Auswahl erfolgte somit nach dem Stand der Wissenschaft.

Für die *Abschreibungen* wurde eine Veränderungsrate von Null angesetzt, da dies dem grundsätzlich handelsrechtlichen Ansatz der Bundesnetzagentur in Bezug auf den Törnqvist-Index entspricht. Eine Alternative zur Ermittlung der kalkulatorischen Abschreibungen des eigenfinanzierten Anteils der Altanlagen auf Basis von Tageswerten wurde in Liebe et al. (2017, S. 87) dargelegt. Dabei wird jedoch vernachlässigt, dass im Übrigen kein Realzins verwendet wird, sondern ein Nominalzins Anwendung findet. Die geforderte Preisentwicklungsberücksichtigung erfolgt also bereits auf der Passivseite, eine doppelte Berücksichtigung wäre demgegenüber nicht sachgerecht. Die Konsistenz mit der sonstigen Vorgehensweise der Bundesnetzagentur spricht damit für die gewählte Methode, die durch die Entscheidung des Bundesgerichtshofs zum Xgen Gas in der dritten Regulierungsperiode bestätigt wurde.¹³⁴

¹³³ Statistisches Bundesamt (2008), S. 37.

¹³⁴ Bundesgerichtshof (2021), Rn.93 ff.

Für die Abbildung der Entwicklung der *Gewerbsteuer* wurden die durchschnittlichen Steuerhebesätze herangezogen. Gegebenenfalls könnte hier zusätzlich auch die Steuermesszahl einfließen, die für die Berechnung der Gewerbesteuer ebenfalls eine Rolle spielt.

In Bezug auf das *Eigenkapital* hat die Bundesnetzagentur folgenden Ansatz gewählt: die von der Behörde in einem eigenständigen Verfahren festgelegten Eigenkapitalzinssätze¹³⁵ wurden als Abbildung der Inputpreisentwicklung angesetzt.¹³⁶ Für die betrachteten Zeiträume (Strom bis 2017 und Gas bis 2016) ergeben sich folgende Werte (vgl. Tabelle 4-4).

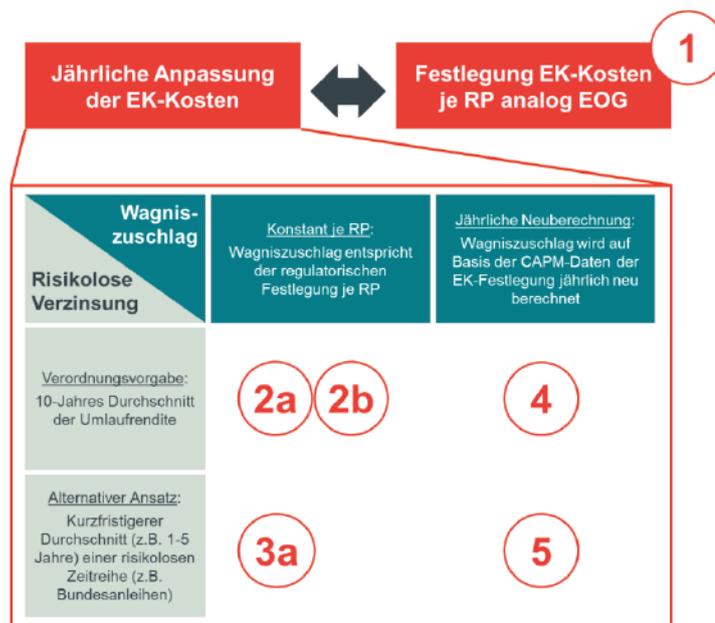
Tabelle 4-4: Entwicklung der Inputpreise in der Position „Eigenkapitalzinsen“

2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
9,21	9,21	9,21	9,29	9,29	9,29	9,29	9,05	9,05	9,05	9,05	9,05

Quelle: Bundesnetzagentur (2018b), S. 41.

Während des Verfahrens wurden verschiedene Alternativen diskutiert, die sich auf die Entwicklung des Eigenkapitals unterschiedlich auswirken. In Abbildung 4-3 sind mögliche Alternativen aufgeführt, die in einem eigenen Gutachten durchgespielt wurden.

Abbildung 4-3: Optionen zur Abbildung der Entwicklung der Eigenkapitalzinsen



Quelle: Frontier Economics (2017), S. 3.

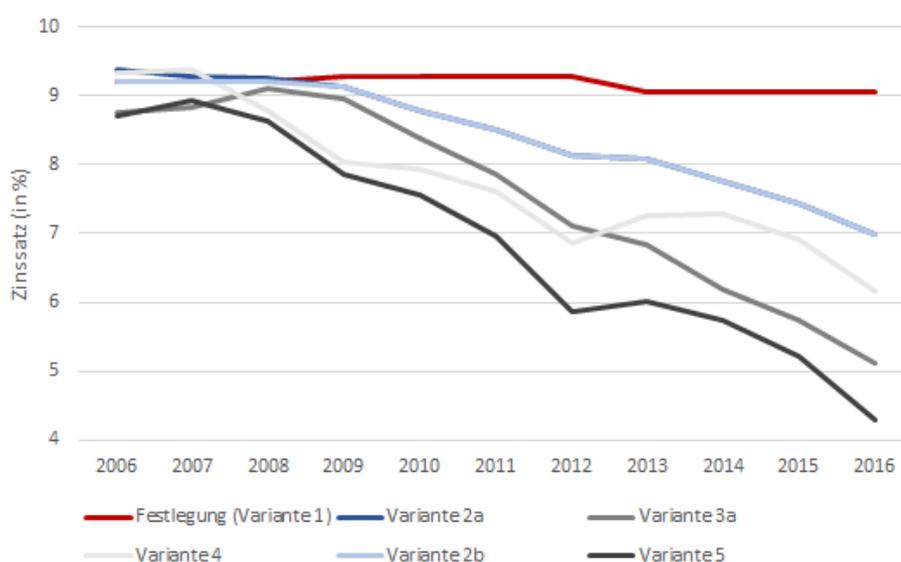
¹³⁵ Vgl. BNetzA (2016a), BNetzA (2016b).

¹³⁶ BNetzA (2018a) S. 33, BNetzA (2018b), S. 41.

Die Unterschiede bestehen in der Festlegung des Wagniszuschlags und der risikolosen Verzinsung. Der Wagniszuschlag kann entweder über die Regulierungsperiode konstant gehalten oder jährlich neu berechnet werden. Die risikolose Verzinsung kann als 10-Jahres-Durchschnitt der Umlaufrendite oder als kurzfristigerer Durchschnitt einer alternativen risikolosen Zeitreihe berechnet werden. In dieser Matrix können sodann unterschiedliche alternative Optionen definiert werden.

Diese Optionen würden sich in einer unterschiedlichen Entwicklung der Eigenkapitalzinsen niederschlagen, wie Abbildung 4-4 zeigt.

Abbildung 4-4: Entwicklung der Eigenkapitalzinsen für verschiedene Optionen (Gas)



Quelle: Eigene Darstellung auf Basis von Frontier Economics (2017), S. 6.

Wie zu sehen ist, ist die von der Bundesnetzagentur gewählte Variante diejenige, bei der die Eigenkapitalzinsen im Zeitablauf am geringsten sinken. Die durchschnittliche Veränderungsrate über den Betrachtungszeitraum beträgt -0,18 %, während sie bei der am stärksten abfallenden Variante 5 einen Wert vom -6,8 % annimmt. Variante 1 wirkt sich also für die Netzbetreiber am günstigsten in Bezug auf den Xgen aus, da der „Preis“ des Eigenkapitals am schwächsten abfällt und die Inputpreise am wenigsten stark sinken.

Die Varianten 2 bis 5 zielen auf eine jährliche Anpassung des Eigenkapitalzinssatzes ab und unterscheiden sich in der Berechnung seiner einzelnen Bestandteile. Diese Bestandteile sind der risikolose Zinssatz und der Wagniszuschlag, welcher sich wiederum aus der Marktrisikoprämie und dem Beta-Risikofaktor zusammensetzt.¹³⁷

¹³⁷ Vgl. im Folgenden: Frontier Economics (2017), S. 4 ff.

Während für Variante 2 der risikofreie Zinssatz jährlich aktualisiert wird, wird für den Wagniszuschlag der für die Regulierungsperiode festgelegte Satz gewählt. In der Variante 2a wird für die Jahre 2006 bis 2008 der Wagniszuschlag der 1. und 2. Regulierungsperiode herangezogen.¹³⁸ In Variante 3a wird in Abweichung von Variante 2a eine alternative Zeitreihe für den risikolosen Zinssatz und eine kürzere Mittelwertbildung verwendet. In Variante 4 wird der 10-Jahresdurchschnitt der Umlaufrendite zuzüglich eines jährlich angepassten Wagniszuschlags berechnet, in Variante 5 ein kurzfristigerer Durchschnitt einer alternativen risikolosen Zeitreihe zuzüglich eines jährlich angepassten Wagniszuschlags.

In der folgenden Tabelle werden die Vor- und Nachteile der verschiedenen Varianten aufgeführt.

Tabelle 4-5: Vor- und Nachteile der Varianten

Variante	Nachteile	Vorteile
Festlegung (Variante 1)	<ul style="list-style-type: none"> Abweichen von kurzfristigen Marktentwicklungen 	<ul style="list-style-type: none"> Robust gegenüber kurzfristigen Verzerrungen Ergebnis eines umfangreichen Prozesses mit entsprechenden Abwägungen Konsistenz mit Regulierungsrahmen Geringer administrativer Aufwand Konservativer Ansatz
Variante 2	<ul style="list-style-type: none"> Kurzfristige Verzerrungen durch Ausreißerjahre möglich 	<ul style="list-style-type: none"> Berücksichtigung der (jährlich) veränderten Risikoeinschätzung durch Kapitalgeber Methodisch nahe an EK-Zins Festlegung Relativ geringer administrativer Aufwand (2a)
Variante 3	<ul style="list-style-type: none"> Abweichung vom Regulierungsrahmen durch Rückgriff auf alternative Zeitreihe Kurzfristige Verzerrungen möglich 	<ul style="list-style-type: none"> Berücksichtigung der (jährlich) veränderten Risikoeinschätzung durch Kapitalgeber Relativ geringer administrativer Aufwand
Variante 4	<ul style="list-style-type: none"> Signifikant hoher Aufwand durch jährliche Berechnung des Beta-Faktors Kurzfristige Verzerrungen möglich 	<ul style="list-style-type: none"> Berücksichtigung der (jährlich) veränderten Risikoeinschätzung durch Kapitalgeber Gute Abbildung aktueller Marktgegebenheiten Methodisch nahe an EK-Zins Festlegung
Variante 5	<ul style="list-style-type: none"> Abweichung vom Regulierungsrahmen durch Rückgriff auf alternative Zeitreihe Kurzfristige Verzerrungen möglich Signifikant hoher Aufwand durch jährliche Berechnung des Beta-Faktors 	<ul style="list-style-type: none"> Berücksichtigung der (jährlich) veränderten Risikoeinschätzung durch Kapitalgeber Beste Abbildung aktueller Marktgegebenheiten

Quelle: Eigene Darstellung auf Basis von Frontier Economics (2017), S. 7 ff.

¹³⁸ In Variante 2b wird ein impliziter Wagniszuschlag aus den in den Jahren 2006 bis 2008 angesetzten EK-Zinsen herangezogen.

Alle aufgezeigten Optionen sind somit mit Vor- und Nachteilen behaftet und entsprechen gleichzeitig dem Stand der Wissenschaft. Die Bundesnetzagentur hat sich mit Variante 1 für einen konservativen Ansatz entschieden.

4.2.5 Gewichtungen

An verschiedenen Stellen in der Berechnung des Törnqvist-Index werden Gewichtungen vorgenommen. Diese werden im Folgenden diskutiert und ihre Vor- und Nachteile aufgezeigt.

Grundsätzlich gewichtet der Törnqvist-Index die eingehenden Berechnungsgrößen mit dem arithmetischen Mittel der Anteile in zwei aufeinander folgenden Jahren (vgl. Abschnitt **Fehler! Verweisquelle konnte nicht gefunden werden.**). Um diese Berechnungsgrößen zu erhalten, werden auch an anderer Stelle Gewichtungen vorgenommen.

Eine Gewichtung erfolgt zunächst für die Berechnung der Inputpreise und des Inputs. Die Gewichte zur Bildung der entsprechenden Indizes entspringen jeweils der GuV-Statistik auf Basis der Datenabfrage für Gas bzw. Strom.

Eine Gewichtung erfolgt auch für die Bildung des Deflators für die Umsatzerlöse, der sich aus den Netzentgelten für Haushalts-, Gewerbe- und Industriekunden zusammensetzt. Hier wurde auf Daten des BDEW bzw. BMWi zurückgegriffen und der durchschnittliche Anteil der Kundengruppen über den Betrachtungszeitraum gemittelt.¹³⁹ Als Alternative hätte eine Berücksichtigung der jährlichen Anteile bei der Berechnung erfolgen können.

Beide Verfahren sind möglich, ohne dass eines der beiden aus wissenschaftlicher Sicht überlegen erscheint. Durch die konstante Gewichtung wurde Konsistenz mit der Vorgehensweise des Statistischen Bundesamtes bei der Berechnung der für die Ermittlung des generellen sektoralen Produktivitätsfaktors verwendeten übrigen Deflatoren (Preisindizes) wie z. B. dem "Index der Erzeugerpreise gewerblicher Produkte" hergestellt. Dieser Index setzt sich aus diversen Subindizes zusammen, die mit unterschiedlichen Gewichten in diesen Preisindex eingehen. Diese Gewichte wurden konstant auf dem Gewicht des Jahres 2010 belassen, eine Anpassung erfolgte hier analog der übrigen Deflatoren nicht.

Auch der Ordnungsgeber wendet im Energiesektor feste Gewichte bei der Zusammensetzung verschiedener Indizes an, etwa bei der Ermittlung von Tagesneuwerten (also sozusagen einer umgekehrten Deflationierung bzw. Inflationierung). In § 6a Abs. 1 der GasNEV heißt es dazu:

„Bei der Ermittlung der Tagesneuwerte nach § 6 Absatz 3 Satz 2 sind folgende Indexreihen des Statistischen Bundesamtes heranzuziehen: [...]

¹³⁹ Vgl. Bundesnetzagentur (2018a), S. 21 und Bundesnetzagentur (2018b), S. 27.

3. für die Anlagengruppen IV.1.1 Stahlleitungen PE ummantelt, IV.1.2 Stahlleitungen kathodisch geschützt und IV.1.3 Stahlleitungen bitumiert, der Anlage 1, die für den Gas-transport mit einem Druck größer als 16 bar ausgelegt sind,

a) die Indexreihe Stahlrohre, Rohrform-, Rohrverschluss- und Rohrverbindungsstücke aus Eisen und Stahl (Statistisches Bundesamt, Fachserie 17, Index der Erzeugerpreise gewerblicher Produkte) mit einem *Anteil von 40 Prozent* und

b) die Indexreihe Ortskanäle, Bauleistungen am Bauwerk (Tiefbau), ohne Umsatzsteuer (Statistisches Bundesamt, Fachserie 17, Preisindizes für die Bauwirtschaft) mit einem *Anteil von 60 Prozent*;

Durch die Mittelwertbildung über die Jahre ist die Bundesnetzagentur bei der Ermittlung eines Deflators für die Umsatzerlöse gewissermaßen einen Kompromiss zwischen beiden Möglichkeiten eingegangen. Dies erkennt zum einen eine sich ändernde Gewichtung über die Zeit an, orientiert sich aber gleichzeitig an der Vorgehensweise des Statistischen Bundesamtes. Der Bundesgerichtshof bemerkt für den Gasbereich hierzu:¹⁴⁰

„Die gegenteilige Auffassung kann nicht auf Berechnungen gestützt werden, nach denen eine jahresscharfe Gewichtung der Verbrauchsanteile zu einem um etwa 0,45 Prozentpunkte geringeren Wert des generellen sektoralen Produktivitätsfaktors führt. Ob und in welchem Umfang jahresscharf bestimmte Verbrauchsanteile sich zugunsten oder zulasten der Netzbetreiber auswirken, ist unerheblich und indiziert weder, dass nur eine jahresscharfe Gewichtung der Verbrauchsanteile dem Stand der Wissenschaft entspräche, noch dass ein solcher Ansatz zur Preisbereinigung der Umsatzerlöse deutlich besser geeignet wäre als der verwendete Index.“

4.2.6 Mittelwertbildung

An verschiedenen Stellen bzw. auf verschiedenen Ebenen werden bei der Berechnung mit Hilfe des Törnqvist-Tools der Bundesnetzagentur für Strom und Gas Mittelwerte gebildet. Diese werden im Folgenden aufgezeigt und mögliche Alternativen diskutiert.

Die Berechnung des Xgen erfolgt auf Basis der etwas umgeformten Gleichung von Bernstein und Sappington (1999, S.11).

$$(4-1) \quad X_{Gen,t} = (\Delta TF_t^{Netz} - \Delta TF_t^{GW}) + (\Delta P_{Input,t}^{GW} - \Delta P_{Input,t}^{Netz}) \\ = (\Delta TF_t^{Netz} - \Delta P_{Input,t}^{Netz}) + \Delta VPI_t.$$

Diese Formel ist allgemein anerkannt und entspricht dem Stand der Wissenschaft.

Die einzelnen Bestandteile der Formel (ΔTF_t^{Netz} , $\Delta P_{Input,t}^{Netz}$ und ΔVPI) werden durch Anwendung des geometrischen Mittels über die Jahre 2006 bis 2016 (Gas) bzw. 2017 (Strom)

¹⁴⁰ Bundesgerichtshof (2021), Rn. 56.

gebildet. Als Alternative bietet sich das arithmetische Mittel an. Tabelle 4-6 zeigt, welche Werte sich jeweils bei den verschiedenen Möglichkeiten der Mittelwertbildung ergeben.

Tabelle 4-6: Mittelwertbildung der Summanden des Xgen

	Geometrisches Mittel	Arithmetisches Mittel
Strom	1,82 %	1,95 %
Gas	0,49 %	1,52%

Quelle: Eigene Darstellung / Berechnung.

Wie zu erkennen ist, würde sich bei einer Anwendung des arithmetischen Mittels (wie erwartet, vgl. Abschnitt 2.3.3) ein höherer Xgen ergeben, was sich zu Ungunsten der Netzbetreiber auswirken würde. Die Anwendung des geometrischen Mittels folgt also dem Gebot der Vorsicht.

Wie in Abschnitt 2.3.3 erläutert, spielt bei Anwendung des geometrischen Mittels die Reihenfolge der Mittelwertbildung bei der Zusammensetzung von Indexreihen (vertikal oder horizontal) keine Rolle.

Auch international findet das geometrische Mittel Anwendung, so z. B. in den Niederlanden. Die dortige Regulierungsbehörde schreibt dazu:¹⁴¹ „ACM stützt die Schätzung der zukünftigen Produktivitätsänderung auf die von den Netzbetreibern selbst bereitgestellten Daten über einen möglichst langen repräsentativen Messzeitraum und unter Verwendung eines geometrischen Mittels (im Folgenden: langfristige Produktivitätsänderung). Damit wird der Ansatz aus der Regulierungsperiode 2017-2021 fortgeführt.“

Die Berechnung der Xgen-Werte für jedes einzelne Jahr und eine anschließende Mittelung wäre ebenfalls eine mögliche Vorgehensweise gewesen.¹⁴² Dies widerspricht allerdings dem Gedanken, dass die Entwicklungen der verschiedenen Bestandteile über die Zeit betrachtet werden. Bei Bildung des arithmetischen Mittels über das gesamte Zeitintervall wäre das Ergebnis dasselbe. Da ja aber, wie in Abschnitt 2.3.3 dargelegt, das geometrische Mittel zum Ansatz kommt, um (Einzel-)Entwicklungen über die Zeit abzubilden, kann auch der gewählte Ansatz angewendet werden.

Die britische Regulierungsbehörde Ofgem basiert ihren Beschluss zur Festlegung der TFP („ongoing efficiency“) auf ähnliche Art und Weise, indem Zeitreihen verschiedener Vergleichsindustrien herangezogen und die Werte für den Betrachtungszeitraum gemittelt werden. Anschließend dienen diese verschiedenen Mittelwerte als Spannweite für die Festlegung des TFP-Wertes.¹⁴³

¹⁴¹ ACM (2021b), Rn. 297.

¹⁴² Demgegenüber hat die BNetzA zunächst die Mittelwerte aller Bestandteile des Xgen (TFP, Inputpreise, VPI) über den Zeitraum ermittelt und diese Werte in die angepasste Formel nach Bernstein / Sappington (1999) eingesetzt.

¹⁴³ Ofgem (2020), S.49.

Bei der Berechnung des Outputindex wurde durch die Bundesnetzagentur auf eine Anwendung der Törnqvist-Formel verzichtet, sondern direkt die Summe der Größen „Umsatzerlöse“, „Bestandsveränderungen“ und „Aktivierte Eigenleistungen“ gebildet, um den Produktionswert zu berechnen. Dies ist aus zwei Gründen nachvollziehbar.

Zum einen existieren im Gasbereich negative Bestandsveränderungen für die Jahre 2013 und 2016, so dass für diese (beiden) Jahre keine Anteile gebildet werden können. Zum andern ist der Anteil der Umsatzerlöse am Produktionswert so hoch, dass eine Anwendung der Törnqvist-Logik den Wert des Xgen im Vergleich zur gewählten Methodik für den Stromsektor nicht verändert. Da die Verhältnisse für den Gasbereich sehr ähnlich ausfallen, kann davon ausgegangen werden, dass auch dort eine direkte Addition der Einzelgrößen das Ergebnis nicht verändert.

Schließlich wurde bei der Berechnung der Inputpreise für das Eigenkapital eine etwas andere Methode der Mittelwertbildung gewählt als für das Fremdkapital. Während für das Eigenkapital ein rollierender 10-Jahres-Mittelwert angesetzt wird (siehe Abschnitt 4.2.4), wurde für das Fremdkapital eine jahresscharfe Betrachtung angewendet.

Stehen Anlagegüter mit unterschiedlichen Nutzungsdauern auf der Aktivseite, so ist normalerweise nicht festzustellen, welche mit kurzfristigem Fremdkapital, welche mit langfristigem Fremdkapital und welche mit Eigenkapital finanziert wurde. Es gilt mithin eher eine pauschalierte Fristenkongruenz (Anlagevermögen und langfristiges Umlaufvermögen soll langfristig, kurzfristiges Umlaufvermögen soll kurzfristig finanziert werden).

Allerdings ist diese Regel theoretisch nicht fundiert und daher umstritten. Denn die Einhaltung der Regeln garantiert nicht die Zahlungsfähigkeit des Unternehmens, bzw. eine Nichtbeachtung führt nicht automatisch zur Insolvenz. Vielmehr kommt es auf die Kreditwürdigkeit des Unternehmens an und somit die Möglichkeit sich am Kapitalmarkt mit frischem Geld einzudecken bzw. auch kurzfristig umzuschulden. Eine Einhaltung der oben beschriebenen Regeln indiziert weder eine besonders gute noch eine besonders schwache Bonität des Unternehmens. Vielmehr muss die Kapitalstruktur so gestaltet sein, dass Kreditgeber die Solvenz des Unternehmens als valide betrachten.¹⁴⁴ Daher kommt es z. B. auch auf andere Kennwerte an, die die Geschäftsentwicklung widerspiegeln (Umsatz, Absatz, Kundenentwicklung etc.). Eine schematische Einhaltung der Regel ist also nicht zwingend, eine jährliche Betrachtung, die ja gerade den Zweck hat, die Situation *aller* Netzbetreiber abzubilden, erscheint demzufolge schlüssig.

Im Hinblick auf die Position der Netzbetreiber als regulierte Unternehmen sind sie auf dem Kapitalmarkt ebenso in einer Wettbewerbsposition wie jedes andere Unternehmen. Sie sind in der Art und Weise ihrer Finanzierung grundsätzlich frei. Es kann daher unterstellt werden, dass sie die optimale Finanzierungsstruktur wählen können, auch was die Fristigkeit betrifft. Die Auswahl der Indexreihen zum Fremdkapitalzins spiegelt neben unterschiedlichen Fristen auch verschiedene Möglichkeiten der Finanzierung und damit verschiedene Risikoprofile wider (Umlaufrenditen inländischer Inhaberschuldverschreibungen / Anleihen der öffentlichen Hand Umlaufrenditen inländischer Inhaberschuldver-

¹⁴⁴ Vgl. z. B. Süchting (1995), S. 490 ff.

schreibungen / Hypothekendarlehen, Umlaufrenditen inländischer Inhaberschuldverschreibungen / Anleihen von Unternehmen (Nicht MFIs)). Die Netzbetreiber unterliegen an dieser Stelle also keinen anderen Regeln als die Unternehmen anderer Branchen bzw. die Gesamtwirtschaft. Der Fremdkapitalmarkt ist derselbe.

Insgesamt erscheint die angewendete Mittelwertbildung für die erläuterten Sachverhalte sachgerecht und nachvollziehbar. Sie entspringen sachlogischen Gründen, internationalen Erfahrungen und praktischen Erwägungen.

4.3 Zwischenfazit

Die bisherige Anwendung des Törnqvist-Index durch die Bundesnetzagentur hat gezeigt, dass es viele Stellschrauben gibt, die einen Einfluss auf das Ergebnis besitzen. Letztlich muss bei allen Berechnungen ein möglichst praktikabler Weg beschritten und verschiedene Alternativen abgewogen werden. Dabei sollen die Methoden stets dem Stand der Wissenschaft entsprechen. Dies ist, wie die Analyse gezeigt hat, in der Berechnung des Törnqvist-Index für die dritte Regulierungsperiode erfolgt.

5 Malmquist-Methode

5.1 Grundansatz

Wie in Abschnitt 2.2.4 eingeführt, entsteht durch den Malmquist-Index ein Maß für die dynamische Effizienzentwicklung über die Zeit. Die Berechnung des Malmquist kann auf der Basis einer Produktions- oder Kostenfunktion gebildet werden:

- Bei einer Produktionsfunktion wird der maximal mögliche Output gemessen, welcher mit einer gegebenen Inputmenge erzielt werden kann;
- eine Kostenfunktion gibt an, welche minimalen Kosten notwendig sind, um einen gegebenen Output bei gegebenen Faktor- bzw. Inputpreisen zu erzielen (bei einer vorgegebenen Produktionsfunktion).

Ein Malmquist-Index, dem eine Kostenfunktion zugrunde liegt, bezeichnen wir im Rahmen dieser Studie als Kostenmalmquist. Der Kostenmalmquist zeigt an, wie stark die Kosten gesenkt werden können, um ein gegebenes Outputniveau Y zu erreichen. Da die Outputs, z. B. Last, Arbeit oder Anzahl der Anschlüsse, vorliegend als exogene Versorgungsaufgabe angenommen werden kann, bietet sich die Berechnung eines Kostenmalmquist-Index an.

Der Kostenmalmquist kann periodenabhängig sowohl im Vergleich zur Referenztechnologie der ersten als auch im Vergleich zur Referenztechnologie der zweiten Periode berechnet werden. Anhand des geometrischen Durchschnitts der beiden Perioden kann ein Kostenmalmquist gebildet werden, welcher von den Perioden unabhängig ist. Wird der Kostenmalmquist anhand der nominellen Kosten (als Inputmengen mal nicht deflatierte Faktorpreise) gebildet, kann der resultierende nominale Kostenmalmquist (MI^{nom}) wie folgt in den Catch-up (CU) und den Frontier-Shift (FS) zwischen den beiden Perioden zerlegt werden:

$$(5-1) \quad MI^{nom} = \left[\frac{E(2,1)}{E(1,1)} \cdot \frac{E(2,2)}{E(1,2)} \right]^{0,5} = \underbrace{\frac{E(2,2)}{E(1,1)}}_{CU^{nom}} \cdot \underbrace{\left[\frac{E(2,1)}{E(2,2)} \cdot \frac{E(1,1)}{E(1,2)} \right]^{0,5}}_{FS^{nom}}$$

Die Terme $E(x,x)$ in Formel (5-1) entsprechen den mittels DEA oder SFA-Methodik berechneten Effizienzwerten (vgl. Kapitel 2.2.4). Die Details zur formalen Herleitung werden in Liebe et al. (2017) genauer ausgeführt.

Der so berechnete nominale Frontier-Shift beinhaltet neben dem technologischen Fortschritt auch die Inputpreisveränderungen des Sektors und bildet somit genau das ab, was vorliegend für den sektoralen Xgen als Abweichung zum VPI relevant ist und berechnet werden muss.

5.2 Anwendung durch die Bundesnetzagentur in der RP3

Da in den Daten der Netzbetreiber keine Informationen zu Inputpreisen enthalten sind, kann der Kostenmalmquist nicht im ursprünglichen Sinne ermittelt werden, sondern nur als „TOTEX-Malmquist“. Für dessen Berechnung sind zwar die Outputmengen bekannt, aber keine Inputmengen und Inputpreise je Netzbetreiber, sondern nur ihr Produkt in Form der Gesamtkosten (TOTEX).

Meyer et al. (2020) führen zwei Bedingungen aus, welche für eine unverzerrte Schätzung des Frontier-Shift durch den TOTEX-Malmquist nötig sind (vgl. Abschnitt 2.1.2). In einem ersten Fall ist der TOTEX-Malmquist-Index unverzerrt, wenn die Effizienzgrenze von Unternehmen gesetzt wird, die entweder in beiden Perioden technisch und allokativ effizient sind oder wenn die Ineffizienzen dieser Grenzünternehmen über die Zeit konstant bleiben.¹⁴⁵ Die Annahme technischer Effizienz ist bei Netzbetreibern, welche Teil der Effizienzgrenze sind, angemessen, da die Anreizregulierung bereits lange in Kraft ist. Auch die Annahme allokativer Effizienz ist plausibel, da in einem Effizienzvergleich mit einem großen Datensatz nur die Netzbetreiber Teil der Effizienzgrenze sind, welche nicht nur technisch, sondern auch allokativ effizient sind. Für eine signifikante Änderung der allokativen Effizienz über die Zeit und über die unterschiedlichen Netzbetreiber, welche die Bedingung für eine solche Verzerrung wäre, gibt es keine Hinweise. Selbst bei Verletzung dieser Annahme ist der TOTEX-Malmquist unverzerrt, wenn das Verhältnis der Inputpreise konstant bleibt. In einem zweiten Fall ist der TOTEX-Malmquist unverzerrt, wenn die Inputpreise konstant sind oder sich bei allen Netzbetreibern im selben Verhältnis verändern. Konstante Inputpreise sind hierbei nicht gegeben. Jedoch ist die Annahme, dass es keine netzbetreiberspezifischen Inputpreisänderungen gibt, angemessen. Durch die bundes- oder weltweite Ausschreibung von Dienstleistungsaufträgen, Auslagerungen etwa von administrativen Funktionen oder der Verlagerung des Hauptsitzes und operativer Tätigkeiten wie bspw. der Netzsteuerung können die Netzbetreiber gegebenenfalls den Inputfaktor Arbeit so optimieren, dass sich keine unterschiedliche Veränderung der Inputpreise zwischen den Netzbetreibern ergibt. Auch die tariflich verhandelten Löhne führen zu einem gewissen Gleichklang regionaler Lohnentwicklungen in Deutschland. Die Verwendung eines TOTEX-Malmquist durch die BNetzA führt somit zu einer unverzerrten Schätzung des Frontier-Shift und hat den Vorteil, dass eine potenzielle Fehlerquelle vermieden wird, weil keine Berechnung von sektoralen Inputpreisänderungen nötig ist.¹⁴⁶

Die Bundesnetzagentur verwendete für den TOTEX-Kostenmalmquist bislang die Netzbetreiberdaten aus den statischen Effizienzvergleichen der Strom- und Gasnetzsektoren aus den ersten drei Regulierungsperioden RP1, RP2 und RP3. Neu kommt die RP4 dazu. Die Basisjahre für den Stromsektor sind 2006, 2011, 2016 und 2021, wobei eine Regulierungsperiode jeweils einen Zeitraum von 5 Jahren umfasst. Für den Gasbereich sind

¹⁴⁵ Meyer et al. (2020), S.95.

¹⁴⁶ Liebe et al. (2017), S.24.

die Basisjahre 2006, 2010, 2015 und 2020, wobei eine Regulierungsperiode einen Zeitraum von 4 (erste Regulierungsperiode) bzw. 5 Jahren (ab der zweiten Regulierungsperiode) umfasst.

Die Gesamtkosten wurden in standardisierter (sTOTEX) und nicht standardisierter Form (TOTEX) jeweils für alle Zeitpunkte und getrennt für den Strom- und Gasbereich verwendet. Die bei den statischen Effizienzvergleichen verwendeten Vergleichs- bzw. Outputparameter sind jedoch sowohl für die Strom- als auch die Gasnetzbetreiber für jede Periode unterschiedlich. Um die Produktivitätsentwicklung über die Zeit messen zu können, ist es notwendig, dass die verwendeten Outputfaktoren in allen betrachteten Perioden vergleichbar sind. Für die Vergleichbarkeit von zwei Regulierungsperioden im dynamischen Effizienzvergleich muss die Modellspezifikation inklusive der darin verwendeten Vergleichsparameter konstant gehalten werden. Je Periodenvergleich kann somit entweder das statische Effizienzvergleichsmodell der früheren oder späteren Regulierungsperiode herangezogen werden. Beim Periodenvergleich RP12 beinhalten die Datensätze der RP1 somit auch die Outputparameter der RP2 und umgekehrt. Äquivalent dazu enthalten beim Periodenvergleich RP23 die Datensätze der RP2 auch die Outputparameter der RP3 und vice versa.¹⁴⁷

Pro Periode resultieren so acht Spezifikationen und Resultate (DEA und SFA jeweils mit Spezifikation der früheren und späteren Regulierungsperiode sowie jeweils für sTOTEX und TOTEX) und es stellt sich die Frage, wie diese Resultate zusammengeführt werden sollen.

Die Bundesnetzagentur verwendete dabei eine Mittelwertbildung, wie sie in Abschnitt 2.3.3 beschrieben ist.¹⁴⁸

Das Vorgehen der Bundesnetzagentur zur Umsetzung des Malmquist-Indexes war nicht unumstritten und Gegenstand von Gerichtsverfahren. Eine Übersicht gibt der nachfolgende Abschnitt 5.3. Ausgewählte Punkte werden später in Abschnitt 5.4 vertieft.

5.3 Streitpunkte aus der RP3 und gerichtliche Beurteilung

Die Anwendung des Malmquist-Ansatzes bei der Ermittlung des generellen sektoralen Produktivitätsfaktors für Betreiber von Gasversorgungsnetzen für die dritte Regulierungsperiode durch die Bundesnetzagentur wurde gerichtlich überprüft und durch den Bundesgerichtshof letztinstanzlich bestätigt (BGH, Beschluss vom 21.01.2021, EnVR 7/20). Betreffend die Festlegung des generellen sektoralen Produktivitätsfaktors für Betreiber von Elektrizitätsversorgungsnetzen für die dritte Regulierungsperiode stellt das Oberlandesgericht Düsseldorf in seiner Entscheidung vom 16.03.2022 fest: „Nach Maßgabe des eingangs dargelegten, eingeschränkten Prüfungsmaßstabs erweist sich auch die von der

¹⁴⁷ Vgl. Liebe et al. (2017), S.44f.

¹⁴⁸ Vgl. auch Liebe et al. (2017), S.65.

Bundesnetzagentur vorgenommene Ermittlung des generellen sektoralen Produktivitätsfaktors nach der Malmquist-Methode als **rechtsfehlerfrei**.“ (OLG Düsseldorf, Beschluss vom 16.03.2022, 3 Kart 637/19 (V), Rn. 201).

Den beiden Beschlüssen folgend, hat die Bundesnetzagentur den generellen sektoralen Produktivitätsfaktor nach der Malmquist-Methode **zutreffend** ermittelt und rechtsfehlerfrei bestimmt:

1. Der methodische Ausgangspunkt ist rechtmäßig.
2. Die Auswahlentscheidungen (u. a. betreffend die Datengrundlage) sind hinreichend transparent und ausreichend begründet.
3. Die angewendeten Berechnungsmethoden (SFA und DEA) sind gerechtfertigt. Eine Überschreitung des methodischen Spielraums ist nicht aufgezeigt. Dies beinhaltet insbesondere die Verwendung von nichtfallenden Skalenerträgen bei der DEA sowie Zeitinteraktionstermen und Halbnormalverteilung des Effizienzterms zur Vermeidung von Konvergenzproblemen bei der SFA.
4. Die praktizierte einfache Mittelwertbildung bei der Ergebnisaggregation anstatt einer Bestabrechnung ist sachgerecht. Rechtmäßig ist auch die Entscheidung, die standardisierten Kosten (sTOTEX) bei der Berechnung zu berücksichtigen.
5. Der auf Basis der Malmquist-Methode ermittelte finale Wert für den generellen sektoralen Produktivitätsfaktor von 0,92 % für Gasnetzbetreiber und von 1,35 % für Stromnetzbetreiber ist nicht zu verwerfen bzw. ist angemessen und genügend robust.

Nachfolgend werden ausgewählte Entscheidungen zu strittigen Punkten näher dargestellt.

5.3.1 DEA: Annahmen zu Skalenerträgen

Die Bundesnetzagentur wendet jeweils die Modellspezifikation der einen Regulierungsperiode auf die jeweils angrenzende Regulierungsperiode an und umgekehrt.

Für die Berechnung des DEA-Malmquist bedeutete dies die Annahme nichtfallender Skalenerträge jeweils dann, wenn die Spezifikation bzw. Vergleichsparameter der ersten und zweiten Regulierungsperiode verwendet wurden (zweimal beim Vergleich der RP1 mit der RP2 sowie einmal beim Vergleich der RP2 mit der RP3). Konstante Skalenerträge werden jedoch für den Vergleich der RP2 und der RP3 verwendet, wenn die Modellspezifikation bzw. Vergleichsparameter der RP3 herangezogen werden, da die ARegV seit der dritten Regulierungsperiode die Annahme konstanter Skalenerträge vorgibt. Während bei konstanten Skalenerträgen alle Netzbetreiber miteinander verglichen werden, werden bei nichtfallenden Skalenerträgen kleine Netzbetreiber nur mit kleinen Netzbetreibern verglichen, große hingegen mit dem gesamten Sample.¹⁴⁹ Die Annahme nichtfallender Skalenerträge wurde zum Schutz kleiner Netzbetreiber für die Anfänge der Anreizregu-

¹⁴⁹ Vgl. auch Tabelle 2-2 auf Seite 24 oder Liebe et al. (2017), S.31.

lierung eingeführt. Die Annahme konstanter Skalenerträge im Energiesektor ist sachgerecht, wenn die Wahl der Betriebsgröße im Ermessen der Netzbetreiber liegt (z. B. durch Zusammenschlüsse). Die Annahme der freien Wählbarkeit der Betriebsgröße sei laut Bundesnetzagentur zu Beginn der Anreizregulierung nicht sachgerecht gewesen, etwa weil Anpassungen der Betriebsgrößen nicht ohne Weiteres umsetzbar gewesen seien.¹⁵⁰

In den Beschwerdeverfahren wurde teilweise die Verwendung konstanter Skalenerträge für alle drei Regulierungsperioden gefordert. Dabei wurde mit dem Xgen als unternehmensübergreifende Vorgabe argumentiert und auf Netzbetreiber verwiesen, welche ihre Betriebsgröße fortlaufend angepasst hätten.¹⁵¹

Die gerichtliche Überprüfung fiel zu Gunsten der Bundesnetzagentur aus, die Verwendung nichtfallender Skalenerträge ist angemessen. Die Argumentation der Beschwerdeführerin, dies entspräche beim dynamischen Effizienzvergleich nicht dem Stand der Wissenschaft, wird als nicht durchgreifend betrachtet.¹⁵²

Die Diskussion der verwendeten Skalenerträge bei der DEA wird nachfolgend in Abschnitt 5.4.5 nochmals aufgenommen.

5.3.2 SFA: „Pooled“ Grundstruktur mit Zeitinteraktionstermen

Wesentliche Einwände betreffend die Grundstruktur mit Zeitinteraktionstermen sind, dass die bei der SFA-Methode gewählte Modellierung mit vollständigen Zeitinteraktionstermen von den für den statischen Effizienzvergleich definierten Modellierungsgrundsätzen der Bundesnetzagentur abweicht, überspezifiziert und unnötig kompliziert ist. Die übermäßige Komplexität lasse die Integration weiterer Aspekte aufgrund von Konvergenzproblemen nicht zu, so dass Aufholeffekte („Catch-up“) nicht mehr berechnet werden könnten.¹⁵³ Zudem erlaube die verwendete Methode keine Separierung der Effizienzgrenzenverschiebung von individuellen Aufholeffekten.¹⁵⁴

Die Bundesnetzagentur rechtfertigt die Verwendung von Zeitinteraktionstermen aufgrund der dynamischen Betrachtung über die Zeit. Während bei der Messung der statischen Effizienz die wesentlichen Parameter diejenigen sind, welche die Effizienz determinieren, sind beim dynamischen Effizienzvergleich diejenigen Parameter wichtig, welche die Veränderung der Effizienz über die Zeit erfassen.¹⁵⁵ Die behauptete Überspezifikation des Modells ist unproblematisch, da für eine unverzerrte Schätzung der Effizienzgrenzenverschiebung die Signifikanz der Regressionskoeffizienten keine Rolle spielt.

Der Bundesgerichtshof hat für den Xgen Gas entschieden, dass die verwendete Methode der wissenschaftlichen Praxis entspreche. Obwohl keine explizite Modellierung des Auf-

¹⁵⁰ Oberlandesgericht Düsseldorf (2022), S.19.

¹⁵¹ Oberlandesgericht Düsseldorf (2022), insbesondere S.8f.

¹⁵² Vgl. Oberlandesgericht Düsseldorf 2022), S. 84f.

¹⁵³ Vgl. Oberlandesgericht Düsseldorf (2022), S.9.

¹⁵⁴ Bundesgerichtshof (2021), Rn. 141.

¹⁵⁵ Liebe et al. (2017), S.35.

holeffekts erfolge, werde eine Trennung von Catch-up und Frontier-Shift über den Ineffizienzterm berücksichtigt.¹⁵⁶ Mit Blick auf den Xgen im Stromsektor hat das Oberlandesgericht Düsseldorf die Verwendung einer sog. „Pooled SFA“ mit Zeitinteraktionstermen nicht beanstandet, da zur Ermittlung des Frontier-Shifts keine explizite Modellierung des Catch-up nötig sei.¹⁵⁷

Das Thema der Grundstruktur der SFA (inklusive Zeitinteraktionstermen) wird nachfolgend in Abschnitt 5.4.5 nochmals aufgenommen.

5.3.3 Modell-Aggregation / Mittelwertbildung

Während sich die Bundesnetzagentur für eine Aggregation gemäß dem in Abschnitt 2.3.3 vorgestellten vierstufigen Prozedere ausspricht, haben die Netzbetreiber eine Bestabrechnung für vorzugswürdig erachtet. Danach sollte analog zum statischen Effizienzvergleich eine „Best-of-four“ Abrechnung aus DEA, SFA, TOTEX und sTOTEX vorgenommen werden. Dabei wird im dritten Schritt für jeden Netzbetreiber individuell das „beste Ergebnis“, also der tiefste Wert dieser Grundspezifikation gewählt, der c.p. zu einem niedrigeren Malmquist-Index führt. Für die Ergebnisse aus den unterschiedlichen Stichproben, Zeitproben und Kostentreiberkombinationen soll jedoch der ungewichtete Mittelwert verwendet werden. Alternativ wurde auch eine „Best-of-four“-Abrechnung mit einem gewichteten Mittelwert aus den unterschiedlichen Stichproben, Zeitproben und Kostentreiberkombinationen oder eine „Best-of-two“ Abrechnung aus TOTEX und sTOTEX vorgeschlagen. Als Begründung zur Verwendung einer Bestabrechnung anstelle eines Mittelwertes wurden hohe Unterschiede zwischen den TOTEX- und sTOTEX-Ergebnissen, zwischen den verschiedenen Methoden (DEA und SFA) sowie bei geringfügigen Änderungen der Daten angeführt. Die Gewichtung des Mittelwerts der einzelnen Stichproben soll anhand der statistischen Unsicherheit erfolgen. Daher wurde vorgeschlagen, diejenigen Ergebnisse stärker zu gewichten, welche ein schmaleres Konfidenzintervall des Frontier-Shifts haben.

Die Bundesnetzagentur hat sich jedoch in der dritten Regulierungsperiode gegen eine solche Gewichtung ausgesprochen, da diese mit einem Informationsverlust verbunden sei. Eine Bestabrechnung aus DEA und SFA lehnt sie ab, da der so ermittelte Durchschnittswert keine ökonomische Aussagekraft hätte, weil das Ergebnis als Kombination aus unterschiedlichen Effizienzgrenzen und Verschiebungen der Effizienzgrenzen bestehe.

Die Entscheidung des Oberlandesgerichts Düsseldorf zum Xgen Strom sieht keine Notwendigkeit zur Verwendung einer Bestabrechnung und bestätigt die Entscheidung der Bundesnetzagentur, eine ungewichtete Mittelwertbildung zu verwenden.¹⁵⁸ Dies deckt sich mit dem früheren Beschluss des Bundesgerichtshofs zum Xgen Gas, wonach eine einfache Mittelwertbildung ohne Bestabrechnung auch bei einer hohen Bandbreite der Ergebnisse der DEA und der SFA eine den wissenschaftlichen Standards genügende und

¹⁵⁶ Bundesgerichtshof (2021), Rn. 141.

¹⁵⁷ Vgl. Oberlandesgericht Düsseldorf (2022), insbesondere S.19.

¹⁵⁸ Vgl. Oberlandesgericht Düsseldorf (2022), insbesondere S.88f.

prinzipiell geeignete Aggregationsmethode im Rahmen des § 9 Abs. 3 S. 1 ARegV darstelle.¹⁵⁹ Ebenso hat der Bundesgerichtshof festgehalten, dass beim Xgen im Vergleich zum individuellen Effizienzwert beim Effizienzvergleich kein individuelles Schutzbedürfnis bestehe.¹⁶⁰

Die Forderung einer Best-of-Abrechnung wird nachfolgend in Abschnitt 5.4.2 noch einmal aufgenommen und erneut verworfen.

5.3.4 Standardisierte Gesamtkosten (sTOTEX)

Überdies wurde gefordert, dass, sofern die Bundesnetzagentur weiterhin die Aggregation durch Mittelwertbildung verwenden sollte, die standardisierten Gesamtkosten (sTOTEX) aus den Berechnungen zu entfernen seien. Bei den sTOTEX werden die Kapitalkosten durch den Annuitätensatz konstant gehalten, im Gegensatz zu den TOTEX, wo die Kapitalkosten durch die lineare Abschreibungsmethode mit dem Alter der Anlage sinken. Als Argument gegen die standardisierten Gesamtkosten wird angeführt, dass die Investitionsgüter, welche bereits teilweise abgeschrieben sind (ältere Netze), bei diesen Netzbetreibern zu einer deutlich stärkeren Verringerung der Eigenkapitalrendite führen. Dies würde sich stärker auf die sTOTEX als auf die TOTEX auswirken und somit den Produktivitätsfaktor deutlich nach oben verzerren.¹⁶¹

Die Bundesnetzagentur argumentiert, dass durch die sTOTEX eine Abmilderung von Verzerrungen aus unterschiedlichen Altersstrukturen der Anlagen sowie abweichender Abschreibungs- und Aktivierungspraktiken der Netzbetreiber erreicht wird, welche die Vergleichbarkeit der Kapitalkosten als Teil der im Effizienzvergleich zu verwendenden Gesamtkosten erhöht.¹⁶²

Das Gericht sah die Verwendung von standardisierten Gesamtkosten in der Berechnung als rechtmäßig an.¹⁶³

Wir teilen diese Entscheidung. Das Thema wird nachfolgend noch einmal kurz aufgegriffen.

5.4 Diskussion

Nachfolgend werden nun ausgewählte Detailfragen näher diskutiert. Die folgenden Kapitel enthalten Ergebnisse für den vorläufigen generellen sektoralen Produktivitätsfaktor Gas der vierten Regulierungsperiode. Die zur Berechnung verwendeten Daten basieren auf dem Datensatz PFG4_Analysen_Daten_VNB und wurden von der BNetzA am 28.02.2023 zur Verfügung gestellt.

¹⁵⁹ Vgl. Bundesgerichtshof (2021), Rn. 122.

¹⁶⁰ Bundesgerichtshof (2021), insbesondere Rn. 121.

¹⁶¹ Oberlandesgericht Düsseldorf (2022), S.97.

¹⁶² Liebe et al. (2017), S.64.

¹⁶³ Vgl. Oberlandesgericht Düsseldorf (2022), insbesondere S.97.

5.4.1 Grundkonzeption

Hinsichtlich der Grundkonzeption, wie der Produktivitätsfortschritt zwischen den Regulierungsperioden gemessen werden soll, ergeben sich zwei Möglichkeiten zur Berechnung des Xgen:

- ein direkter Vergleich der Perioden 1 und 4 oder
- ein stufenweiser Vergleich der Perioden 1 und 2, der Perioden 2 und 3 sowie der Perioden 3 und 4 mit anschließender Mittelung dieser drei Frontier-Shifts.

Eine stufenweise Verrechnung hat den Vorteil, dass sie zu größeren Datensätzen führt, da gewisse Betriebe zwar in mehreren Effizienzvergleichen aktiv waren, nicht aber im ersten oder letzten Effizienzvergleich. So können beispielsweise Netzbetreiber, welche in den ersten drei Perioden existieren, aber in der vierten nicht mehr, immer noch für die Frontier-Shifts 12 und 23 verwendet werden, während sie bei einem direkten Vergleich nicht in die Berechnung einfließen würden. Auch hinsichtlich der Outputparameter vergrößert eine stufenweise Berechnung die Datenbasis. Gewisse Outputvariablen der aktuellsten Periode können womöglich nicht mehr für die erste Periode ermittelt werden und müssten aus der Analyse entfernt werden. Bei einer stufenweisen Verrechnung führt die größere Datenbasis mit mehr Beobachtungen und Outputvariablen zu einer robusteren Schätzung und kann die Dynamik des Sektors besser abbilden, da sie die Produktivitätsfortschritte in mehrere Teilperioden unterteilt.

Eine solche **stufenweise Berechnung** des Xgen wurde von der Bundesnetzagentur bisher wie folgt durchgeführt:

- stufenweise Periodenvergleiche (RP12, 23, in Analogie neu auch 34) mit Mittelwert quer und geometrischem Mittel längs über die Zeit;
- Verwendung der beiden Methoden SFA und DEA;
- Anwendung dieser auf die zwei Kostenarten TOTEX und sTOTEX;
- Je Periodenvergleich, Methode und Kostenart jeweils einmal Spezifikation der früheren und einmal der späteren Regulierungsperiode.

Die vierstufige Berechnung des generellen sektoralen Produktivitätsfaktors erfolgt für **zwei Fälle**: Fall 1 enthält alle Netzbetreiber, in Fall 2 werden nur Netzbetreiber berücksichtigt, deren Fläche sich zwischen den Perioden nicht stark verändert hat (Entfernung aller Netzbetreiber, bei denen sich die versorgte Fläche und das Konzessionsgebiets um mehr als 10 % verändert hat).

Die resultierenden acht Werte je Periodenvergleich führen bei nunmehr drei Periodenvergleichen (12, 23 und neu 34) für jeweils 2 Fälle zu insgesamt 48 Werten, die zunächst je Fall gemäß dem Vorgehen in Abschnitt 2.3.3 in eine eindeutige Schätzung für den generellen sektoralen Produktivitätsfortschritt überführt werden:

- Stufe 1: Da für jeden Periodenvergleich die Spezifikation des früheren und des späteren statischen Effizienzvergleichs gerechnet wird, werden die beiden Werte

mit dem geometrischen Mittel verrechnet. Somit erhalten jeweils beide Spezifikationen das gleiche Gewicht.

- Stufe 2: Eine Verrechnung der Methoden SFA und DEA durch das arithmetische Mittel erlaubt es, die Stärken der beiden Modelle gleichgewichtet zu kombinieren.
- Stufe 3: Da hinsichtlich der Kostenarten keine der anderen überlegen ist, führt auch hier eine Verrechnung durch das arithmetische Mittel zu einer robusteren Schätzung des Frontier-Shift.
- Stufe 4: Da es sich bei den 3 Periodenvergleichen um eine Zeitreihe handelt, ist eine Verrechnung durch das geometrische Mittel angemessen. Somit werden die Mittelwerte aus den 8 Modellen je Regulierungsperiode zu einem das gesamte Zeitintervall überspannenden durchschnittlichen Produktivitätsfortschritt verrechnet.

Der dadurch ermittelte Wert für Fall 1 und Fall 2 wird durch das arithmetische Mittel zum generellen sektoralen Produktivitätsfaktor aggregiert.

Die obigen Punkte (stufenweises Vorgehen mit zugehöriger Mittelwertbildung) erachten wir grundsätzlich als weiterhin richtig und empfehlen demnach, dies auch in der laufenden Festsetzung des Xgen für die vierte Regulierungsperiode unverändert anzuwenden.

5.4.2 Mittelwertbildung oder „Best-of“

Das Anliegen einer „Best-of“-Verrechnung in der dritten Stufe des Vorgehens der Bundesnetzagentur bei der Mittelwertbildung zum Malmquist-Index (vgl. Abschnitt 2.3.3) entstammt aus den relativen Effizienzvergleichen der Bundesnetzagentur, wo jeweils der beste Effizienzwert eines Netzbetreibers aus vier Schätzungen („Best-of-four“) herangezogen wird, sofern dieser über 60 % liegt. Ansonsten wird der Mindesteffizienzwert angewendet.

Vorliegend wird jedoch weniger ein „Best-of“ gefordert, sondern vielmehr ein „Worst-of“, indem von den vier Grundspezifikationen (SFA/DEA je für TOTEX/sTOTEX) jeweils der tiefste errechnete Wert herangezogen werden soll, also von allen Spezifikationen die schlechteste errechnete Produktivitätsverbesserung.

Aus analytischer Sicht ist ein solches Vorgehen nicht angezeigt, da sich Effizienzvergleich und Abschätzung des Xgen methodisch fundamental unterscheiden:

Xind – individuelle relative Effizienz:

- Beim Effizienzvergleich wird die **individuelle** Effizienz (Xind) eines Netzbetreibers in einem relativen Effizienzvergleich ermittelt, also **relativ** zur jeweiligen Best Practice bzw. ermittelten Effizienzgrenze, die von wenigen Netzbetreibern aufgespannt wird. Aus dem relativen Vergleich folgt die Vergleichbarkeit als zentrales Anliegen. Liegt beim Vergleich z. B. ein Datenfehler vor, kann dies für das betreffende Unternehmen zu einer zu hohen oder zu tiefen gemessenen Effizienz führen. Liegt ein Unternehmen bei der DEA ohne Datenfehler nicht auf der Effizienzgrenze, hat ein Datenfehler, der dessen Kosten erhöht, keine Auswirkungen auf

die übrigen Unternehmen. Im umgekehrten Fall (also bei einem Datenfehler, der die Kosten senkt) können jedoch, falls das Unternehmen so zu einem Peer wird, die Effizienzwerte vieler anderer Unternehmen nach unten verzerrt werden.¹⁶⁴ Ein analoger Zusammenhang gilt bei Datenfehlern auf der Outputseite. Dies bedeutet, dass sich Datenfehler bei der DEA im Erwartungswert zu Lasten der Netzbetreiber insgesamt auswirken.

- Gleichzeitig ist die Verteilung der Ineffizienz **asymmetrisch**, beispielsweise wird sie bei der SFA i.d.R. als halbnormal oder exponentiell angenommen. Schätzfehler können nicht zu einer negativen Ineffizienz führen (Werte über 1), sondern nur Effizienz mindern oder an den Wert 1 annähern. Dies bedeutet, dass der Erwartungswert für einen theoretisch angenommen effizienten Netzbetreiber kleiner als 1 liegt, da Daten- oder Schätzfehler zu einer Effizienz tiefer als 1 führen können, nie aber höher als 1.
- Diese beiden Punkte stützen das von der Bundesnetzagentur angewendete **Vorsichtsprinzip** bzw. die in der ARegV in § 12 Abs. 4a angelegte **Bestabrechnung**.

Xgen - sektoraler Fortschritt:

- Der Xgen wird nicht individuell, sondern sektoral für alle Netzbetreiber gleichermaßen festgelegt. Schätzfehler führen zu einer Über- oder Unterschätzung des wahren Wertes. Eine Unterschätzung bedeutet eine Begünstigung der Netzbetreiber zu Lasten der Netznutzer, bei einer Überschätzung ist es umgekehrt.
- Die Schätzfehler beim Malmquist sind **symmetrisch verteilt**: Es kann gleichermaßen zu einer Über- und Unterschätzung kommen (Werte über und unter 1 möglich).
- Ziel bei der Bestimmung des Xgen muss daher eine **möglichst akkurate Schätzung** sein, welche keine Seite (Netzbetreiber/Netznutzer) systematisch bevorteilt. Die Anwendung eines Vorsichtsprinzips im Sinne der Netzbetreiber würde zu einer systematischen Benachteiligung der Netznutzer führen.

Im Übrigen ist davon auszugehen, dass bei einer „Best-of“ Verrechnung jeweils Schätzwerte unterschiedlicher Spezifikationen resultieren, die dann anschließend in Schritt 4 miteinander vermengt werden. Beispielsweise könnte beim Vergleich RP12 der Wert einer DEA mit TOTEX am tiefsten sein, beim Vergleich der RP23 aber der Wert einer SFA mit sTOTEX. Die Vereinigung völliger unterschiedlicher Grundspezifikationen anhand des geometrischen Mittels in Schritt 4 stellt eine zusätzliche mögliche Fehlerquelle dar bzw. dürfte zu systematisch verzerrten Resultaten führen.

Aufgrund dieser zusätzlichen Überlegungen ist die Empfehlung, an der Mittelwertbildung, wie sie in Abschnitt 2.3.3 zusammengefasst ist, festzuhalten.

¹⁶⁴ Wäre das Unternehmen bereits vor dem Datenfehler ein Peer in der DEA gewesen, gilt hingegen ein grundsätzlich symmetrischer Zusammenhang.

5.4.3 Vergleichsparameter: Auswahl

Im Rahmen des Verfahrens zur Bemessung des Xgen der RP3 wurde der Ansatz der Bundesnetzagentur von den Netzbetreibern dahingehend gerügt, dass sie keine dynamische Kostentreiberanalyse vorgenommen habe. Konkret stellt sich die Frage, ob die Vergleichsparameter aus den vergangenen Effizienzvergleichen auch geeignet sind, Produktivitätsänderungen in der jeweiligen Netzbranche zu messen. Bei der Auswahl der Vergleichsparameter können bei der Anwendung des Malmquist-Indexes grundsätzlich folgende Vorgehensweisen gewählt werden:

- Verwendung des Modells der vorangegangenen Regulierungsperiode auch mit Daten der neuen Regulierungsperiode und umgekehrt (d. h. die jeweils gleiche Modellspezifikation, aber mit veränderten Daten; hierzu sind ggf. aus den Altdaten neue Parameter zu bilden);
- Wahl einer neuen Modellspezifikation (z. B. angepasste Auswahl an Vergleichsparametern) auf Grundlage einer neuen, dynamischen Kostentreiberanalyse.

Keine Option ist der direkte Vergleich der je Regulierungsperiode gewählten Effizienzvergleichsmodelle (unterschiedliche Parameter mit unterschiedlichen Daten), da so keine Aussage zur Veränderung zwischen den Perioden getroffen werden kann.

Bei der Wahl, welcher der beiden obigen Ansätze umgesetzt werden soll, spielen zum einen analytische Überlegungen eine Rolle und zum anderen müssen Vergleichsparameter gewählt werden, die auch praktisch verwendbar sind. Z. B. muss die SFA einer vorangegangenen Regulierungsperiode mit den neuen Daten nicht unbedingt konvergieren oder gewisse Vergleichsparameter können nur in einer Regulierungsperiode überhaupt gebildet werden, weil zwischenzeitlich die Datenabfragen geändert wurden. Daneben spielen auch technische Voraussetzungen der beiden Methoden eine Rolle. Z. B. sollte die Zahl der Vergleichsparameter nicht zu groß sein, um eine Überspezifikation zu vermeiden, und mit Blick auf die DEA müssen die Vergleichsparameter skalierbar sein. Gleichzeitig ist zu beachten, dass die bisherigen Modelle für sich genommen für die jeweilige Regulierungsperiode „optimal“ sind, also die statischen Begebenheiten des jeweiligen Basisjahrs sehr gut abbilden. Demgegenüber führt eine neue Modellfindung zu einem offenen Ausgang mit entsprechenden Unabwägbarkeiten für alle Beteiligten.

Für die vorzunehmende Wahl wird in einem ersten Schritt qualitativ untersucht, ob die bestehenden Vergleichsparameter der Effizienzvergleichsmodelle der bisherigen Regulierungsperioden potenziell Produktivitätsveränderungen über die Zeit abzubilden vermögen („dynamische Eignung“). In einem zweiten Schritt wird geprüft, ob die internationale Praxis bei der Anwendung von Malmquist-Indizes in den betreffenden Sektoren auf andere, geeignetere Vergleichsparameter hinweist.

Bislang wurden in den statischen Effizienzvergleichen die in Tabelle 5-1 zusammengefassten finalen Modellspezifikationen verwendet. Im Wesentlichen sind dies Netzlängen, Anschlüsse und/oder Zählpunkte, Höchstlasten sowie in den ersten beiden Regulierungsperioden die Fläche. Beim Strom wurde zudem die dezentrale Erzeugungsleistung verwendet.

Tabelle 5-1: Bislang in Effizienzvergleichen verwendete Vergleichsparameter

EVS1	EVS2	EVS3	EVS4
Anschlüsse Umspannwerke 5 Leitungs- und Kabel- längen 2 Höchstlasten 1 Dezentrale Erzeu- gungsleistung Fläche	Anschlüsse Zählpunkte 5 Leitungs- und Kabel- längen 2 Höchstlasten 1 Dezentrale Erzeu- gungsleistung Fläche	Zählpunkte 4 Leitungs- und Ka- bellängen 2 Höchstlasten 2 Dezentrale Erzeu- gungsleistung	In Erarbeitung (Herbst 2023)
EVG1	EVG2	EVG3	EVG4
2 Ausspeisepunkte Rohrvolumen 2 Netzlängen 2 Höchstlasten 2 Bevölkerungs-zahlen Fläche	3 Ausspeisepunkte Anzahl Messstellen Rohrvolumen Netzlänge Höchstlast Fläche Bodenklasse	Ausspeisepunkte Anzahl Messstellen Rohrvolumen Höchstlast Bodenklasse	Ausspeisepunkte Anzahl Messstellen Rohrvolumen Höchstlast Bodenklasse
EFG1	EFG2	EFG3	EFG4
Transportmoment Ein- und Ausspeise- punkte Fläche	Rohrvolumen Ein- und Ausspeise- punkte Fläche (yArea)	Rohrvolumen Verdichterleistung Ein- und Ausspeise- punkte Fläche (yArea_all)	Rohrvolumen Verdichterleistung Ausspeisepunkte Fläche (yArea)

Für die grundsätzliche dynamische Eignung wird für ausgewählte exogene Schocks und mögliche Maßnahmen der Netzbetreiber analytisch untersucht, inwieweit sich dies für einen Netzbetreiber über zwei Regulierungsperioden in seinen Aufwand- und Vergleichsparametern abbildet und nachgelagert bei der Xgen-Berechnung mit dem Malmquist-Index ceteris paribus auswirkt. Ausgehend von dieser Analyse kann gefolgert werden, welche Parameter welche grundsätzliche Eignung für eine dynamische Frontier-Shift-Analyse aufweisen, bzw. inwieweit die statischen Modelle der RP3 und RP4 angewendet werden können.

Tabelle 5-2 zeigt das Resultat der vorgenommenen Analyse des ersten Schritts. Sämtliche unterstellten Ereignisse, die zwischen den beiden Regulierungsperioden angenommen werden (aufgeführt in der ersten Spalte), führen, soweit abschätzbar, zu einer Veränderung des Xgen in die erwartete Richtung, da die wesentlichen Vergleichsparameter für Strom und Gas in den jeweiligen statischen Effizienzvergleichen berücksichtigt sind.

Etwas weniger eindeutig ist der Effekt im Falle der stromseitigen Energiewende, für die Parameter teils erst ab der dritten (Rückspeisung) bzw. vierten Regulierungsperiode (Anschlüsse/Zählpunkte mit Blick auf Elektromobilität) erhoben wurden. Soweit die angeschlossene dezentrale Erzeugungsleistung in einem proportionalen Zusammenhang zu den hiermit verbundenen Netzkosten ist, ist die Energiewende für den errechneten Xgen neutral. Liegen bei der Integration sinkende oder steigende Skalenerträge vor, entsteht ein Effekt in die eine oder andere Richtung. Gerade dieses Beispiel unterstützt im Übrigen

das Vorgehen der Bundesnetzagentur, jeweils zwei aneinander liegende Regulierungsperioden miteinander zu vergleichen und nicht etwa die RP1 mit der RP4. Da neue Trends zwar jeweils in den Datenabfragen berücksichtigt wurden, jedoch nicht alle davon bereits in der RP1 vorhersehbar waren, können für einen Direktvergleich wichtige Vergleichsparameter in früheren Jahren fehlen. In die Zukunft gedacht, wird sich diese Entwicklung tendenziell verstärken.

Tabelle 5-2: Auswirkungen von ausgewählten Ereignissen

Ereignis/Maßnahme Netzbetreiber	Wirkung auf Inputs (TOTEX, sTOTEX)	Wirkung auf Outputs	Effekt auf Xgen	Fazit
Einsatz einer effizienten neuen Technologie, z. B. Automatisierung in der Zentrale	Reduktion (stärker reduzierender Effekt auf OPEX als erhöhender Effekt auf CAPEX)	-	Positiv	✓
Höhere Lohnkostensteigerung im Vergleich zur Gesamtwirtschaft	Höher	-	Negativ	(✓) ¹⁶⁵
Erhöhung aller Inputpreise im Umfang des VPI	Höher	-	Neutral	✓
Günstigere Leitungsverlegung / günstigerer Unterhalt	Tiefer	-	Positiv	✓
Outsourcing	Neutral	Neutral	Neutral	✓
Erhöhung der Bevölkerung (Anschlüsse)	-	Höher (mehr Anschlüsse, Last, Arbeit, Anschlussleitungen, Leitungen)	Positiv bei Skaleneffekten	✓
Energiewende stromseitig	Höher	Höher (höhere angeschlossene dezentrale Erzeugung, Leistung, Rückspeisung, Einspeisepunkte, ggf. mehr Leitungen)	(Neutral)	(✓)
Energiewende gassseitig	I.d.R. nicht Teil der Aufwandparameter (Biogas, Investmaßnahmen)	I.d.R. Biogas herausgerechnet bei Anschlüssen, Leitungen usw.	Neutral	✓
Mengenrückgänge Gas	Tiefere OPEX (unterproportional)	Tiefer (niedrigere Last, Arbeit)	Negativ	(✓)

Quelle: Eigene Darstellung.

In Österreich, dessen Regulierungspraxis am ähnlichsten ist (vgl. Kapitel 2.1.3.3), werden ähnliche, wenn auch stärker endogene Vergleichsparameter verwendet wie in Deutschland. Für die aktuelle Regulierungsperiode im Stromsektor spielten sowohl für

¹⁶⁵ Problematische Anreizwirkung, wenn höhere Löhne ex post mit c.p. tieferem Xgen belohnt werden.

den Xgen als auch für den Effizienzvergleich die gewichtete transformierte Netzanschlussdichte (Modellnetzlänge), unterteilt nach Netzebene, die Netzhöchstlasten und die Zählpunkte eine Rolle.¹⁶⁶ Für die aktuelle Regulierungsperiode Gas wurden beim Effizienzvergleich die Netzlängen, die Zählpunkte und die verrechnete Leistung von Industrie und großem Gewerbe berücksichtigt. Beim Xgen wurde die Summe der Netzlängen über beide Ebenen sowie die Summe der Zähler über beide Ebenen verwendet.¹⁶⁷

Gestützt auf die obigen Überlegungen sehen wir keinen Grund, vom bisherigen Ansatz abzuweichen, jeweils zwei aneinander liegende Regulierungsperioden zu vergleichen und dabei die jeweiligen Modellspezifikationen der statischen Effizienzvergleiche auf die jeweils andere Periode zu übertragen, um den Malmquist-Index zu bestimmen. Die Parameter der statischen Effizienzvergleiche vermögen, wie analytisch gezeigt, Änderungen in der Produktivität grundsätzlich abzubilden. Darüber hinaus gibt es keine Hinweise aus anderen Regulierungsregimen, dass weitere Parameter relevant sein könnten, die in den vorliegenden Datensätzen nicht abgebildet sind.

5.4.4 Vergleichsparameter: Fläche

Die Fläche wird verschiedentlich in den Effizienzvergleichen Strom und Gas verwendet.

In dem Zusammenhang ist die Rechtsprechung des Bundesgerichtshofes zum EVG2, insbesondere hinsichtlich der Schätzung der versorgten Fläche für die ehemaligen regionalen Fernleitungsnetzbetreiber, relevant (BGH, Beschluss vom 12.06.2018, EnVR 54/17).

Der Bundesgerichtshof hält fest, dass die Auswahl der Vergleichsparameter für den Effizienzvergleich nicht zu beanstanden sei. Insbesondere sei es ermessensfehlerfrei, dass die Parameter "Ausgespeiste Jahreshöchstlast" und "Ausspeisepunkte > 16 bar" einbezogen worden seien. Eine Ermittlung von Teileffizienzen für einzelne Druckstufen sei nicht geboten gewesen.

Mit Blick auf die versorgte Fläche habe die Bundesnetzagentur für die ehemaligen regionalen Fernleitungsnetzbetreiber rechtsfehlerfrei geschätzt, weil diese Unternehmen zum Zeitpunkt der Datenerhebung noch nicht als Verteilernetzbetreiber einzustufen gewesen seien und die für diesen Parameter einschlägigen Informationen deshalb nicht hätten ermittelt werden können. Die stattdessen angestellte Schätzung auf der Grundlage der Flächen aller Gemeinden, durch die Leitungen dieser Unternehmen verliefen, entspreche der Vorgehensweise bei den anderen Netzbetreibern. Hierbei hatte die Bundesnetzagentur die Schlüssel aller Gebiete herangezogen, durch die Leitungen der betreffenden Netzbetreiber verlaufen und in denen zu deren Netz gehörende Anlagen belegen sind.

Damit sei gemäß dem Bundesgerichtshof ein angemessener Vergleich mit Netzbetreibern, die innerhalb eines Konzessionsgebiets ein flächenhaftes Verteilernetz betreiben,

¹⁶⁶ E-Control (2018), S. 35.

¹⁶⁷ E-Control (2018), S. 24f.

nicht gewährleistet. In diesem Zusammenhang könne dahingestellt bleiben, ob ein Gebiet schon dann durch ein Netz versorgt wird, wenn darin eine Leitung verläuft oder eine zum Betrieb eingesetzte Anlage gelegen ist, oder ob erforderlich ist, dass aus dem Netz Gas in das örtliche Verteilernetz eingespeist wird. Auch hinsichtlich von Gebieten, bei denen diese Voraussetzung vorliegt, sei es systemwidrig, in gleicher Weise auf die gesamte Fläche abzustellen wie bei örtlichen Verteilernetzen. Bei örtlichen Verteilernetzen sei die versorgte Fläche als potenzieller Kostenfaktor insbesondere deshalb relevant, weil sich das Netz typischerweise über erhebliche Teile dieser Fläche erstrecken muss, damit die Aufgabe der örtlichen Verteilung erfüllt werden kann. Für Netze, die vor allem dazu dienen, Gas über weitere Entfernungen zu örtlichen Verteilernetzen zu transportieren, besteht ein solcher Zusammenhang typischerweise nicht.

Die versorgte Fläche wurde zunächst im Rahmen des Effizienzvergleichs der dritten Regulierungsperiode analog verwendet, jedoch später nach Veröffentlichung der Entscheidung des Bundesgerichtshofs angepasst (vier Varianten von neuen Flächenparametern gebildet für acht betroffene Unternehmen, vgl. z. B. Tabelle 5 im zugehörigen Ergebnisbericht von Frontier Economics).

Dies bedeutet vorliegend, dass, die bestehenden Flächenparameter des EVG2 für die betreffenden fünf Unternehmen der RP2 nicht gesetzeskonform sind. Soll der Flächenparameter in die Berechnung des Frontier-Shift einbezogen werden, sind für die RP2 entweder die entsprechenden fünf Unternehmen aus dem Datensatz der RP2 bzw. RP3 zu entfernen oder die Parameter nachträglich anzupassen.

Zur Anpassung des Flächenparameters der betreffenden Netzbetreiber bestehen folgende Varianten:

- Übernahme der Parameterwerte aus der RP3 auf die RP2 (so keine Netzübergänge).
- Anwendung von 0-Werten analog der Argumentation des Bundesgerichtshofs, da keine Versorgung in die Fläche erfolgt. In Analogie folgt dieser Ansatz dem EFG4, bei dem eine „vertikale Fläche“ gebildet wird, welche sich von der horizontalen Versorgungsaufgabe abgrenzen soll, für die beim Parameter kein Flächenwert zugeschrieben wird. Beim EFG4 wurde die vertikale Fläche geprüft und schließlich verworfen einerseits aus statistischen Gründen (Multikollinearität) und andererseits aus konzeptionellen Gründen (resultierende Flächenwerte bei FNB mit sowohl horizontaler als auch vertikaler Versorgung).
- Verwendung einer der vier neugebildeten Flächenparameter.

Soweit es sich um ausschließliche Regionalnetzbetreiber ohne eigene vertikale Versorgungsaufgabe handelt, kann aus konzeptioneller Sicht dem Bundesgerichtshof gefolgt werden, d. h. es können Flächenwerte von Null angesetzt werden. Soweit eine vertikale Versorgung an Kunden in der NS-Ebene erfolgt, können die entsprechenden NS-Flächen berücksichtigt werden.

Werden alle Netzbetreiber mit Nullwerten bei der versorgten Fläche zur Berechnung des Frontier-Shift RP23 mit Modell RP2 entfernt, hätte dies eine geringfügig positive Auswirkung auf den resultierenden nominalen Frontier-Shift gegenüber der Anwendung von Nullwerten:

- Bei der DEA führt sie zu einem leicht höheren Frontier-Shift von durchschnittlich 0.9874 anstatt 0.9861.
- Bei der SFA resultiert ein leichter Anstieg des durchschnittlichen Frontier-Shift mit Modellen der RP2 von 1.0336 auf 1.0345.
- Insgesamt erhöht sich der aggregierte Frontier-Shift p.a. für die RP23 geringfügig von 0.1936 auf 0.2065.

Wir empfehlen, an der Anwendung von Nullwerten festzuhalten, soweit bei diesen Netzbetreibern keine Versorgung an eigene Endkunden in die Fläche erfolgt.

5.4.5 Umsetzung DEA: Annahme zu den Skalenerträgen

Der Berechnung der Effizienzwerte in der DEA können unterschiedliche Annahmen zu Skalenerträgen zugrunde liegen (vgl. Abschnitt 2.2.6). In der Anwendung der BNetzA stellt sich die Frage zwischen einer Kombination aus der Verwendung von nichtfallenden und konstanten Skalenerträgen sowie der durchgehenden Verwendung von konstanten Skalenerträgen. Bei nichtfallenden Skalenerträgen wird angenommen, dass die Netzbetreiber ihre Betriebsgröße nicht frei wählen können und es Effizienzvorteile für größere Netzbetreiber gibt. Dabei werden kleine Netzbetreiber nur mit kleinen Netzbetreibern verglichen, große hingegen mit allen, wodurch insbesondere kleinere Netzbetreiber geschützt sind. Die Annahme konstanter Skalenerträge erlaubt den Vergleich eines Netzbetreibers mit allen anderen und basiert auf der freien Wählbarkeit der Betriebsgröße sowie fehlender Effizienzvorteile großer Netzbetreiber.

Es ist plausibel, dass die freie Wählbarkeit der Betriebsgröße (z. B. durch Zusammenschlüsse) und zu Beginn der Anreizregulierung noch nicht gegeben war und größere Netzbetreiber noch über Effizienzvorteile verfügten. Mit fortschreitender Dauer der Anreizregulierung sollte die freie Wählbarkeit der Betriebsgröße für alle Netzbetreiber zutreffen, womit aus regulatorischer Sicht der Bedarf einer Korrektur von Größeneffekten entfällt. Die neun Jahre der ersten beiden Regulierungsperioden sind eine nachvollziehbare Grenze für das Eintreten dieser Bedingungen. Somit sollte auch zur Berechnung des DEA-Malmquist, äquivalent zum statischen Effizienzvergleich, bei Modellspezifikationen der ersten beiden Perioden nichtfallende Skalenerträge und in den anschließenden Perioden konstante Skalenerträge verwendet werden.

Vorliegend folgen wir erneut dem Grundsatz, die Modellspezifikation der einen Regulierungsperiode jeweils auf die angrenzende anzuwenden. D. h. die DEA basierend auf Modellspezifikationen der RP3 und 4 werden jeweils mit konstanten Skalenerträgen gerechnet, die Modellspezifikationen der RP1 und 2 jeweils mit nichtfallenden Skalenerträgen.

5.4.6 Umsetzung SFA: Grundstruktur

Die Berechnung der SFA kann entweder anhand von gepoolten Daten oder Paneldaten erfolgen. Ersteres stellt den Ansatz der Bundesnetzagentur dar.

Paneldatenmodelle stellen einen Bezug zwischen verschiedenen Beobachtungen eines Unternehmens im Zeitverlauf her und können damit individuelle Effekte eines Unternehmens über die Zeit explizit berücksichtigen. Für Paneldatenmodelle werden Beobachtungen der Unternehmen über mehrere Perioden benötigt. Ansätze mit nur zwei Perioden führen normalerweise zu Verzerrungen und unzuverlässigen Resultaten. Da zur Berechnung des Xgen ein Einzelvergleich der Perioden mit anschließender Mittelung stattfindet, können jeweils nur zwei Perioden verwendet werden, wobei aber auch drei oder vier Perioden das Problem nicht wesentlich lindern würden. Im Gegensatz zu gepoolten Daten sind somit weniger verlässliche Aussagen möglich.¹⁶⁸

Bislang hat die Bundesnetzagentur daher gepoolte Daten verwendet. Durch das Poolen der Daten beider Regulierungsperioden zu einem Datensatz und die Anwendung der gleichen Modellspezifikationen kann eine gemeinsame Frontier geschätzt werden. Durch die Verwendung einer Dummy-Variable für die spätere Periode ist somit auch die Veränderung der Kosten über die Zeit berücksichtigt, mit welcher der Frontier-Shift berechnet werden kann.¹⁶⁹ Dieser Frontier-Shift kann jedoch auch andere systematische Veränderungen zwischen den beiden Regulierungsperioden enthalten, welche nicht im Modell abgebildet werden. Die zusätzlich verwendeten Zeitinteraktionsterme ermöglichen, dass sich die Kostenauswirkungen der Kostentreiber über die Zeit verändern und einen Kreuzvergleich der Netzbetreiber mit der Effizienzgrenze der jeweils anderen Periode zulassen. Die Zeitinteraktionsterme sind daher für die Ermittlung der Verschiebung der Effizienzgrenze erforderlich, auch wenn diese statistisch nicht signifikant sind. Durch die Verwendung der Zeitinteraktionsterme kann der Frontier-Shift somit möglichst unverzerrt geschätzt werden. Dabei erfolgt zwar keine explizite Modellierung des Aufholeffekts, allerdings wird über den Ineffizienzterm eine Trennung von Catch-up und Frontier-Shift berücksichtigt.

Ergänzend zum Modell der Bundesnetzagentur kann auch ein Modell mit einer zeitvariablen Ineffizienzverteilung aufbauend auf Battese Coelli (1995), wie es von einer Beschwerdeführerin bei der RP3 Strom verlangt wurde, verwendet werden. Dieses Modell basiert auf der Annahme, dass sich die Verteilung der Ineffizienzen über die Zeit ändert. Eine unterschiedliche Verteilung der Effizienzwerte könnte beispielsweise auftreten, wenn die Aufholeffekte vor allem bei ineffizienteren VNB stattfinden. Die Verteilung der Ineffizienzwerte wird für jede Zeitperiode separat geschätzt, wobei die Varianz der Verteilung des Ineffizienzterms durch eine Konstante (wie im Modell der Bundesnetzagentur) sowie zusätzlich einen zeitabhängigen Teil geschätzt wird. In den vergangenen Regulierungsperioden wies dieses Modell jedoch eine stärkere Differenz zu den Ergebnissen der Malmquist-DEA auf, welche einen individuellen Catch-up jedes VNB beinhalten, als das

¹⁶⁸ Vgl. Agrell et al. (2014), insbesondere S. 35.

¹⁶⁹ Vgl. Liebe et al. (2017), S. 53.

von der Bundesnetzagentur verwendete Modell. Der zusätzlich geschätzte Zeitterm der Ineffizienzverteilung erhöht die Anzahl der zu schätzenden Parameter und kann dadurch zu Konvergenzproblemen oder einer ungenaueren Schätzung führen. Da es in den Daten keine Hinweise gibt, dass sich die Verteilung der Ineffizienzen innerhalb der Regulierungsperioden signifikant ändert, empfiehlt es sich, auf diesen zusätzlichen Parameter zu verzichten. Ein allgemeiner Catch-up wird im Übrigen schon implizit im Modell der Bundesnetzagentur abgebildet, indem die VNB in der späteren Periode tiefere Ineffizienzen (geschätzt durch den Ineffizienzterm u) aufweisen. Für jede Beobachtung im gepoolten Datensatz wird ein Ineffizienzwert geschätzt. Hat ein VNB in der späteren Periode einen tieferen Ineffizienzwert, so impliziert dies, dass ein positiver Catch-up stattgefunden hat.

Die in der dritten Regulierungsperiode Strom im Vergleich zu anderen Modellierungsmethoden gemessene geringere Differenz der Bundesnetzagentur-Herangehensweise zu den DEA-Ergebnissen, welche eine explizite Separierung in Frontier-Shift und Catch-up beinhalten, spricht gemäß dem Oberlandesgericht Düsseldorf für eine Pooled SFA mit Zeitinteraktionstermen und einen Verzicht auf eine periodenabhängige Ineffizienzverteilung nach Battese Coelli (1995).¹⁷⁰

Dies untermauert erneut der Vergleich des Xgen für eine SFA mit und ohne Zeitinteraktionsterme sowie mit einem Modell mit einer zeitvariablen Ineffizienzverteilung und DEA mit den Daten für den PFG4 VNB Gas.

Tabelle 5-3 zeigt die Werte für den generellen Produktivitätsfaktor Gas mit unterschiedlichen Methoden und Spezifikationen. Der PF VNB wurde gemäß der in Kapitel 5.4.1 beschriebenen Grundkonzeption berechnet, mit der einzigen Ausnahme, dass keine Verrechnung der beiden Methoden (DEA und SFA) vorgenommen wurde. Der separat berechnete Xgen für Fall 1 und Fall 2 wurde zu einem PF für die VNB aggregiert. Für die SFA wurde neben der von der Bundesnetzagentur verwendeten Spezifikation mit Zeitinteraktionstermen auch dieselbe Spezifikation ohne Zeitinteraktionsterme sowie eine alternative Spezifikation mit einer zeitvariablen Ineffizienzverteilung nach Battese und Coelli (1995), jeweils unter Annahme einer Halbnormalverteilungsannahme der Ineffizienz, gerechnet.

¹⁷⁰ Vgl. Oberlandesgericht Düsseldorf (2022), S. 19.

Tabelle 5-3: Vergleich der Resultate für PFG4 VNB mit alternativen SFA-Spezifikationen

	PF Fall 1 VNB	PF Fall 2 VNB	PFG4 VNB
Verwendete SFA Spezifikation der Bundesnetzagentur mit Zeitinteraktionstermen	1,2998 %	1,2861 %	1,2929 %
Bisher verwendete SFA Spezifikation ohne Zeitinteraktionsterme	1,2746 %	1,2641 %	1,2693 %
Spezifikation mit variabler Ineffizienzverteilung nach Battese Coelli (1995)	0,6797 %	0,5638 %	0,6217 %
Von der Bundesnetzagentur verwendete DEA Spezifikation	1,2606 %	0,9523 %	1,1065 %

Datengrundlage vom 28.02.2023

Der nach der Spezifikation der Bundesnetzagentur berechnete generelle Produktivitätsfaktor der vierten Periode mit der SFA beträgt 1,2929 % und ist somit leicht höher als bei der DEA mit 1,1065 %. Unterschiede zeigen sich hinsichtlich der beiden Fälle¹⁷¹: beim Fall 1 sind die Resultate praktisch identisch (DEA 1,2606 % und SFA 1,2998 %), bei Fall 2 ist der PF bei der SFA mit 1,2861 % höher als bei der DEA mit 0,9523 %. Im Vergleich zur von der Bundesnetzagentur verwendeten Spezifikation zeigt eine SFA ohne Zeitinteraktionsterme einen leicht tieferen Wert von 1,2693 %. Dies zeigt, dass die Zeitinteraktionsterme zwar für einen korrekten Kreuzvergleich der Netzbetreiber mit der Effizienzgrenze der anderen Periode nötig sind, aber nur einen minimalen Einfluss auf den Xgen haben.

Die Spezifikation mit einer zeitvariablen Ineffizienzverteilung zeigt hingegen erheblich tiefere Werte. Insbesondere die starke Differenz zur DEA, welche eine individuelle Berechnung des Catch-up ermöglicht, deutet darauf hin, dass die zusätzliche Berücksichtigung der Zeitflexibilität der Ineffizienzverteilung die Präzision des geschätzten generellen sektoralen Produktivitätsfaktors vermindert. Die mehrheitlich übereinstimmenden Werte des Xgen bei SFA und DEA sprechen hingegen dafür, dass der Catch-up auch in der SFA ausreichend berücksichtigt wird.

5.4.7 Umsetzung SFA: Annahme zur Verteilung der Ineffizienz

Dem Ineffizienzterm der SFA lagen in den bisherigen Regulierungsperioden aufgrund von Konvergenzaspekten unterschiedliche Verteilungsannahmen zugrunde. So wurde a priori eine Exponentialverteilung der Ineffizienz angenommen, wobei bei Konvergenzproblemen mit diesen Modellen auf eine Halbnormalverteilung umgestellt wurde. Dies führte dazu, dass bei Strom durchgängig in allen Fällen die Halbnormalverteilung angewendet wurde, während bei Gas in der Hälfte der Fälle eine Exponentialverteilung angenommen wurde, wobei dies sowohl für den Frontier-Shift RP12 als auch für RP23 gilt.

¹⁷¹ Fall 1 mit allen VNB und Fall 2 ohne diejenigen VNB, bei denen sich die versorgte Fläche und die Konzessionsfläche um mehr als 10% verändert haben.

Es gibt jedoch weder analytisch noch empirisch Hinweise, dass die exponentielle Verteilung oder die Halbnormalverteilung der jeweils anderen vorzuziehen ist. Auch in der Literatur, wie beispielsweise bei Aigner et al. (1977) mit Halbnormalverteilung und Meeusen et al. (1977) mit Exponentialverteilung gibt es keinen Konsens zur Wahl des Ineffizienzterms. Die exponentielle Verteilung führt in den statischen Effizienzvergleichen der Bundesnetzagentur c.p. zu leicht höheren Effizienzwerten, ist aber ansonsten gleichwertig. Aufgrund dieser höheren Werte wurde von der Branche in den statischen Effizienzvergleichen die Verwendung der Exponentialverteilung gefordert und von der Bundesnetzagentur infolge des Vorsichtsprinzips seit der RP2 auch eingesetzt.

Da die Ermittlung des Xgen jedoch über einen *Frontiervergleich* geschieht, wird der Effekt etwas höherer/geringerer Effizienz durch Differenzenbildung weitgehend eliminiert: Es werden bei Exponentialverteilung also jeweils zwei Effizienzgrenzen verglichen, die mit etwas höheren Effizienzwerten einhergehen, als dies bei Halbnormalverteilung der Fall ist. Beim Vergleich wird der Leveffekt beim Vorgehen der RP3 vollständig eliminiert, da der Frontier-Shift nur durch die Zeitdummy sowie die Zeitinteraktionsterme geschätzt wird, welche unabhängig vom Ineffizienzterm u sind.¹⁷²

Wichtig ist jedoch, dass die Frontiers, die verglichen werden, jeweils mit der gleichen Ineffizienzverteilungsannahme geschätzt wurden. Ansonsten kann eine Verzerrung des Frontier-Shifts in die eine oder andere Richtung entstehen. Wird beispielsweise die frühere Periode mit Halbnormalverteilung geschätzt und die spätere mit Exponentialverteilung, führt dies zu Verzerrungen.

Dieses analytische Resultat wird durch eine empirische Analyse der vergangenen Regulierungsperioden gestützt: Die Resultate der bisherigen Regulierungsperioden zeigen keinen statistisch signifikanten Einfluss der Verteilungsannahme auf den Xgen. So ist für jedes Modell der Unterschied zwischen der Halbnormalverteilung und der Exponentialverteilung insignifikant und auch über alle Modelle gesehen, ist der Xgen nur minimal höher bei der Exponentialverteilung. Tabelle 5-4 zeigt beispielhaft, wie gering die Abweichungen der Werte für den Frontier-Shift p.a. im Gassektor zwischen den beiden Ineffizienzverteilungsannahmen sind. Die Differenz des Frontier-Shift zwischen Halbnormalverteilung und Exponentialverteilung ist in jedem Modell kleiner als 0,007 Prozentpunkte. Da je nach Modell die Berechnungen mit halbnormal- oder exponentialverteilten Ineffizienztermen höhere Werte haben, ist die Differenz nach Aggregation der Zeitperioden sogar noch tiefer. Allerdings gab es bei der Berechnung der SFA mit Exponentialverteilung nicht unerwartet vereinzelt Konvergenzprobleme, insbesondere bei der 3. Regulierungsperiode Gas. Bei der Halbnormalverteilung traten diese Konvergenzprobleme nicht auf.

¹⁷² Vgl. hierzu Formeln 4-2 und 4-3 in Liebe et al. (2017), S. 52 und 53.

Tabelle 5-4: Vergleich der Resultate für den Frontier-Shift p.a. Gas mit Halbnormal- und Exponentialverteilungsannahme des Ineffizienzterms der SFA (Datengrundlage 28.02.2023)

	RP12 Fall 1	RP12 Fall 2	RP23 Fall 1	RP23 Fall 2	RP34 Fall 1	RP34 Fall 2
Halbnormalverteilung	-0,5121	-0,4033	0,4031	-0,0159	-0,2083	-0,3259
Exponentialverteilung	-0,5168	-0,4091	0,4035 ¹⁷³	-0,0090	-0,2087	-0,3243
Unterschied HN-Exp	0,0047	-0,0058	-0,0004	0,0069	0,0004	-0,0016

Demgegenüber verhindert die *durchgängige* Verwendung der Halbnormalverteilung anstelle der Exponentialverteilung Konvergenzprobleme ohne die Einbuße prädiktiver Aussagekraft und ohne signifikante Änderung bzw. Verzerrung des Xgen. Als Konsequenz ist die durchgängige Verwendung der Halbnormalverteilung als Annahme zur Ineffizienzverteilung bei der SFA zu bevorzugen.

Die Wahl der Verteilungsannahme der Ineffizienz bei der SFA sollte im Sinne der Kohärenz auch durch den entsprechenden F-Test für das Dominanzkriteriums der DEA berücksichtigt werden (vgl. Kapitel 5.4.7).

5.4.8 Umsetzung DEA und SFA: Behandlung von Ausreißern

5.4.8.1 Gesetzliche Grundlagen

Die Behandlung von Ausreißern beim statischen Effizienzvergleich ist in der ARegV geregelt (vgl. ARegV Anlage 3 zu § 12). Dabei gelten in der DEA diejenigen Beobachtungen als Ausreißer, welche für einen überwiegenden Teil des Datensatzes als Effizienzmaßstab gelten würden. „Dabei ist die mittlere Effizienz aller Netzbetreiber einschließlich der potenziellen Ausreißer mit der mittleren Effizienz der Netzbetreiber zu vergleichen, die sich bei Ausschluss der potenziellen Ausreißer ergeben würde. Der dabei festgestellte Unterschied ist mit einer Vertrauenswahrscheinlichkeit von mindestens 95 Prozent zu identifizieren.“¹⁷⁴ Nach Entfernung dieser dominanten Beobachtungen soll eine Supereffizienzanalyse durchgeführt werden, bei welcher die Ausreißer entfernt werden, deren Effizienzwerte den oberen Quartilswert um mehr als den 1,5-fachen Quartilsabstand, definiert als Spannweite der zentralen 50 % des Datensatzes, übersteigen. Bei der SFA gilt eine Beobachtung als Ausreißer, wenn sie die Lage der Regressionsgerade in starkem Maß beeinflusst. Als mögliche Methoden werden Cooks Distance, DFBETAS, DFFITS,

¹⁷³ Fehlende Konvergenz bei vier Modellen, darum wurde in diesem Fällen mit Halbnormalverteilung gerechnet.

¹⁷⁴ Vgl. ARegV Anlage 3 zu § 12.

Covariance-Ratio oder Robuste Regression erwähnt. In den statischen Effizienzvergleichen ist dabei stets das Cooks Distance Kriterium angewendet worden.

5.4.8.2 Anwendung in vergangenen Regulierungsperioden

Da für die Malmquist-Berechnungen des generellen sektoralen Produktivitätsfortschritts dieselben Methoden wie beim statischen Effizienzvergleich verwendet werden (DEA und SFA), sollte, wenn möglich und so sich diese in den statischen Effizienzvergleichen bewährt haben, auch bei den Ausreißeranalysen auf dieselben Kriterien wie bei den statischen Effizienzvergleichen zurückgegriffen werden.

Behandlung von festgestellten Ausreißern

Auch wenn die gleichen Methoden zur Ausreißerbereinigung angewendet werden, unterscheidet sich die Ausreißeranalyse im dynamischen Effizienzvergleich zum Statischen, da beim dynamischen Effizienzvergleich Daten für jeweils zwei Perioden in die Analyse einfließen. Für den Frontier-Shift 34 sind dies beispielsweise die Daten der dritten Periode und der vierten Periode. Es stellt sich somit die Frage, in welchen Perioden die Ausreißer identifiziert werden (in beiden oder jeweils nur einer) und wie diese im Datensatz bereinigt werden sollen. Es gibt folgende zwei Methoden, um die Ausreißerbereinigung im dynamischen Effizienzvergleich vorzunehmen:

- **Trimming** – stets in beiden Perioden entfernen: Diese Methode ist in der wissenschaftlichen Praxis weit verbreitet und kann sowohl bei der DEA als auch bei der SFA angewendet werden. Dabei wird ein als Ausreißer klassifizierter Netzbetreiber komplett aus dem Datensatz dieses Modells entfernt, unabhängig davon, ob er in beiden oder nur einer der beiden Perioden ein Ausreißer ist.
- **Winsorizing** – stets in beiden Perioden modifiziert beibehalten: Die Methode kann nur bei der DEA angewendet werden, wobei die Supereffizienzwerte benutzt werden, um einem als Ausreißer klassifizierten Netzbetreiber künstlich einen Effizienzwert zuzuweisen. Dies führt dazu, dass Netzbetreiber sowohl mit den regulären Beobachtungen als auch mit Ausreißern mit angepassten Werten im Datensatz erhalten bleiben. Dabei kann entweder allen „übereffizienten“ Ausreißern ein Effizienzwert von 100 % zugeschrieben werden, was dazu führt, dass diese Beobachtungen zwar auf der Frontier liegen, aber nicht die Effizienzgrenze für diese Periode aufspannen. Als zweite Möglichkeit können die Kosten der Ausreißer gemäß ihres Supereffizienzwertes hochskaliert werden, wodurch sie nicht mehr maßgeblich für die Lage der Effizienzgrenze sind.

In den bisherigen Regulierungsperioden wurde stets ein Trimming vorgenommen. Dies ist auch in der wissenschaftlichen Literatur weitverbreitet und hat gegenüber den beiden Arten von Winsorizing mehrere Vorteile:

- Das Vorgehen ist konsistent zu den statischen Effizienzvergleichen, da auch dort alle Ausreißer aus dem Datensatz entfernt werden und der Effizienzvergleich (DEA und

SFA) nur auf die verbleibenden Beobachtungen angewendet wird. Den Ausreißern wird dabei außerhalb der Effizienzanalyse ein Wert von 1 zugewiesen

- Eine konsistente Ausreißerbereinigung zwischen den beiden Methoden DEA und SFA kann nur bei Trimming erreicht werden. Da Winsorizing zur Berechnung von Ersatzwerten für Ausreißer auf das Supereffizienzkriterium zurückgreift, kann dieses bei der DEA angewendet werden, ist aber bei der SFA nicht möglich.
- Beim Trimming findet keine Veränderung der Ausgangsdaten statt. Beim Winsorizing fließen hingegen künstlich manipulierte Daten in die Effizienzanalyse ein, welche bei den Netzbetreibern nie so gemessen wurden.

DEA: Umsetzung Dominanztest

Zur Identifizierung dominanter Netzbetreiber, welche für einen überwiegenden Teil der Netzbetreiber als Effizienzmaßstab gelten würden und die durchschnittliche Effizienz der Netzbetreiber signifikant verändern, wird ein F-Test verwendet. Die berechneten Kosteneffizienzwerte können dabei mit der Halbnormal- oder Exponentialverteilungsannahme der Ineffizienzen getestet werden. In den bisherigen Regulierungsperioden Strom und Gas wurde stets der Banker Test¹⁷⁵ angewendet mit Annahme einer zugrundeliegenden Halbnormalverteilung der Ineffizienz. Die Teststatistik für die Nullhypothese, dass die Ineffizienzverteilung mit und ohne den getesteten Netzbetreiber $i \neq j$ gleich ist, wird berechnet als

$$(5-2) \sum_{j=1}^{N_1} \left(\frac{(\theta_{j1}-1)^2}{N_1} \right) / \sum_{j=1}^{N_2} \left(\frac{(\theta_{j2}-1)^2}{N_2} \right) \quad F \text{ verteilt mit } (N_1, N_2) \text{ Freiheitsgraden}$$

Wobei θ_{j1} die Effizienz von Netzbetreiber j in einer DEA ohne Netzbetreiber i , θ_{j2} die Effizienz von Netzbetreiber j in einer DEA mit Netzbetreiber i und N_1 sowie N_2 die Anzahl Beobachtungen ohne Netzbetreiber i sind. Soweit die Halbnormalverteilungsannahme beim Ineffizienzterm der SFA angenommen wird (vgl. Abschnitt 5.4.6), sind die beiden Methoden auch hinsichtlich der Verteilungsannahme konsistent. Die Ausreißeranalyse anhand des Supereffizienzkriteriums wurde in den bisherigen dynamischen Effizienzvergleichen anhand der Vorgaben aus der ARegV durchgeführt.

SFA: Cooks Distance

Die Ausreißer bei der Berechnung der SFA wurden anhand des Cooks Distance Kriteriums ermittelt. Als Schwellenwert wurde dafür seit der 2. Regulierungsperiode Strom auf $4/(n-p-1)$ festgelegt. Eine Wilson-Analyse hat identifiziert, dass dieser Schwellenwert zuverlässigere Resultate liefert als der vorgängig verwendete Wert von 1 und somit diesem gegenüber zu bevorzugen sei (vgl. Swiss Economics und Sumicsid 2014a, S.55).

¹⁷⁵ Banker und Chang (1995). S 41.

5.4.8.3 Einschätzung alternativer Ausreißermethoden

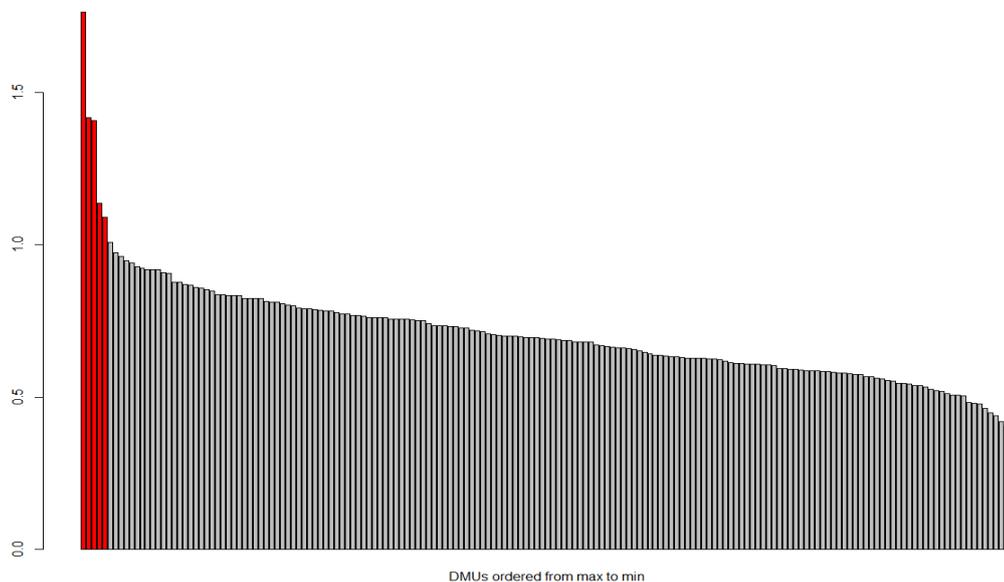
Masking

In den Beschwerdeverfahren zum Xgen in der dritten Regulierungsperiode wurde bei den bestehenden Ausreißermethoden auf sogenannte „Masking Effects“ verwiesen. Diese „Masking Effects“ bedeuten, dass gewisse Ausreißer so extreme Werte haben, dass sie andere potenzielle Ausreißer verdecken. Analoge Forderungen wurden jeweils auch bei den statischen Effizienzvergleichen aufgestellt und verworfen. Zur Identifikation von maskierten Ausreißern wird eine mehrstufige Ausreißeranalyse gefordert. Dabei sollen in drei Stufen eine Dominanz- und Supereffizienzanalyse in der DEA sowie eine Cook Distance Analyse in der SFA durchgeführt werden. Maskierte Ausreißer der ersten Stufe könnten so nach Entfernung der identifizierten Ausreißer der ersten Stufe sichtbar werden.

Eine solche mehrstufige Ausreißerbereinigung widerspricht den Vorgaben aus der ARegV für den statischen Effizienzvergleich. Ein eindeutiger numerischer Wert für den Einfluss auf die Lage der Regressionsgerade, wie von der ARegV gefordert, könnte so nicht berechnet werden. Auch ist das Cook's Distance Kriterium inkl. Schwellenwert für eine einmalige Anwendung eines fixen Schwellenwerts entwickelt worden, vgl. Banker and Chang (2006). Insofern erscheint das Supereffizienz-Kriterium bzw. dessen Schwellenwert auch im vorliegenden Kontext sachgerecht, um Unternehmen als Ausreißer zu identifizieren.

Abbildung 5-1 zeigt die abnehmend sortierten Supereffizienzwerte des Modells der RP3 sTotex Fall 1. Es ist ersichtlich, dass der Test auch vorliegend zu einem plausiblen Ergebnis führt. Die Netzbetreiber mit deutlich höheren Supereffizienzwerten werden alle als Ausreißer identifiziert (rot markiert), während die Supereffizienzwerte der nicht als Ausreißer klassifizierten Netzbetreiber keine Sprünge mehr aufweisen und einer konsistenten Verteilung folgen. Insofern erscheint das Supereffizienz-Kriterium bzw. dessen Schwellenwert auch im vorliegenden Kontext sachgerecht, um Unternehmen als Ausreißer zu identifizieren.

Abbildung 5-1: Geordnete Supereffizienzwerte Modell RP3 Fall 1 Totex



Quelle: Eigene Abbildung

Dynamische Ausreißer

Ein weiterer Diskussionspunkt aus den Beschwerdeverfahren zum Xgen für die dritte Regulierungsperiode ist die Entfernung dynamischer Ausreißer. In einer dynamischen Ausreißeranalyse würden die Beobachtungen, die innerhalb der untersuchten Regulierungsperiode ein extremes Wachstum in einer Outputvariable (positiv oder negativ) haben, entfernt. Als Kriterien für extreme Wachstumswerte genannt werden das Überschreiten des dritten Quartils um mehr als das 1.5-fache des Interquartilabstandes, das Unterschreiten des ersten Quartils um mehr als das 1.5-fache des Interquartilabstandes, eine Entfernung von mehr als zwei Standardabweichungen zum Mittelwert sowie das generelle Entfernen der Beobachtungen der 5 % höchsten und 5 % tiefsten Wachstumsraten genannt.

Eine solche Bereinigung wäre indes fehlerhaft, da dadurch alle Netzbetreiber mit hohen und tiefen Wachstumswerten entfernt würden, unabhängig davon, ob diese Werte extrem oder außergewöhnlich sind. Die Beobachtungen würden auch unabhängig davon entfernt, inwieweit die Wachstumsraten der Normalverteilung entsprechen. Die Entfernung korrekter Datenpunkte würde das Resultat verfälschen, da potenziell aussagekräftige Datenpunkte mit hohen oder tiefen Werten entfernt würden.

Zudem ist bei Netzbetreibern mit sehr hohen positiven oder negativen Wachstumsraten die Wahrscheinlichkeit groß, dass diese in einer der beiden Perioden eine so außerge-

wöhnliche Kosten-Kostentreiberkombination haben, dass sie bereits durch die vorhandenen Ausreißermethoden (Dominanz- und Supereffizienzkriterium bei der DEA sowie Cook's-Distancekriterium bei der SFA) identifiziert werden.

Schließlich werden Netzbetreiber, welche im dynamischen Kontext außergewöhnliche Veränderungen des Versorgungs- und Konzessionsgebiets haben, durch Fall 2 (vgl. Abschnitt 5.4.1) bereits als potenzielle Ausreißer berücksichtigt. Für Fall 2 werden alle Netzbetreiber entfernt, welche zwischen den beiden Perioden eine Veränderung des Versorgungs- und Konzessionsgebiets von mehr als 10 % hatten. Diese Flächenveränderungen zeigen strukturelle Netzveränderungen an, welche eine Vergleichbarkeit zwischen den beiden Perioden erschweren können. Mit dem Vorteil der verbesserten Vergleichbarkeit der beiden Perioden geht jedoch auch der Nachteil eines verkleinerten Datensatzes, sowie des Verlusts potenziell aussagekräftiger Beobachtungen einher. Da keiner der beiden Fälle klar zu präferieren ist, stellt die separate Berechnung der Fälle 1 und 2 mit anschließender Mittelung eine robuste Lösung dar.

6 Fazit

Die in dieser Studie diskutierten Verfahren zur Bestimmung des generellen sektoralen Produktivitätsfaktors für die vierte Regulierungsperiode Strom und Gas (nach Törnqvist und Malmquist) entsprechen dem Stand der Wissenschaft. Dies ergibt sich sowohl aus der entsprechenden wissenschaftlichen Literatur, der Rechtsprechung des Oberlandesgerichts Düsseldorf und des Bundesgerichtshofs sowie der internationalen Regulierungspraxis.

Eine wesentliche Stärke des Malmquist-Indexes ist dabei die Berechnung des Frontier-Shifts auf Basis von Unternehmensdaten, während in den Törnqvist-Index neben diesen auch Daten bzw. Zeitreihen aus öffentlich verfügbaren Quellen einfließen. Der Törnqvist-Index hat den Vorteil, dass die Methode einfacher nachzuvollziehen ist, jedoch Annahmen bezüglich technischer, allokativer und Skaleneffizienz trifft. Bei der Berechnung des Malmquist-Indexes können dagegen Frontier-Shift und Catch-up berücksichtigt werden.

In der Überprüfung der bisherigen methodischen Ausgestaltung zur Berechnung des generellen sektoralen Produktivitätsfaktors (in der dritten Regulierungsperiode) zeigt sich, dass gewisse Freiheitsgrade bestehen (Beobachtungszeitraum, Mittelwertbildung etc.). Es wurde aufgezeigt, mit welchen Verfahren die Berechnungen validiert werden können.

Für die Törnqvist-Methode wurde gezeigt, dass es verschiedene Stellschrauben gibt, die einen Einfluss auf das Ergebnis besitzen. Letztlich muss bei allen Berechnungen das Ziel eines robusten Ergebnisses verfolgt und verschiedene Alternativen abgewogen werden, aber gleichzeitig ein möglichst praktikabler Weg beschritten werden.

Für den Malmquist-Index wurde die Methodik der dritten Regulierungsperiode entsprechend an die vierte Regulierungsperiode angepasst. Auch hier wurden verschiedene Aspekte (z. B. Mittelwertbildung, Auswahl der Vergleichsparameter) nochmals diskutiert, um die Berechnungen für verschiedene Spezifikationen mittels DEA und SFA durchzuführen. Dabei wurde auch auf den Umgang mit Ausreißern eingegangen.

Auch die vom Bundesgerichtshof bestätigte Vorgehensweise zur Ermittlung der gesamtwirtschaftlichen Bestandteile nach § 9 ARegV wurde nochmals einer kritischen Würdigung unterzogen. Dabei konnten keine Anhaltspunkte identifiziert werden, die einer Anwendung der Residualmethode im Wege stehen.

Im Ergebnis stehen die jeweils berechneten Ergebnisse vor dem Hintergrund der jeweils durchgeführten Diskussionen somit auf einem validen Fundament. Die auf Basis dieses validen methodischen Fundaments ermittelten Ergebnisse für den generellen sektoralen Produktivitätsfaktor sind als robust einzustufen.

7 Literatur

- ACM [Autoriteit Consument en Markt](2015): Gewijzigd methodebesluit regionale netbeheerders elektriciteit, 2014 – 2016, Besluit van de Autoriteit Consument en Markt als bedoeld in artikel 41, eerste lid, van de Elektriciteitswet 1998.
- ACM [Autoriteit Consument en Markt](2021a): Methodebesluit regionale netbeheerders elektriciteit 2022-2026, Besluit van de Autoriteit Consument & Markt als bedoeld in artikel 41, eerste lid, van de Elektriciteitswet 1998.
- ACM [Autoriteit Consument en Markt](2021b): Methodebesluit regionale netbeheerders gas 2022-2026, Besluit van de Autoriteit Consument & Markt als bedoeld in artikel 81, eerste lid, van de Gaswet.
- ACM [Autoriteit Consument en Markt] (2021c): Bijlage 1 bij het methodebesluit regionale netbeheerders gas 2022- 2026, Uitwerking van de methode in rekenkundige formules.
- Agrell et al. (2014): Effizienzvergleich für Verteilernetzbetreiber Strom 2013, Ergebnisdokumentation und Schlussbericht.
- Agrell, P., Bogetoft, P., Koller, M., Trinkner, U. (2014): Effizienzverbesserungen VNB Strom. Veränderung der Effizienz der Verteilernetzbetreiber von der ersten zur zweiten Regulierungsperiode. Bericht im Auftrag der Bundesnetzagentur (Bundesnetzagentur), September 2014
- Aigner, D.J., Lovell, C.A.K., Schmidt, P. (1977): Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6: 21–37.
- Ajayi, V., Anaya, K., M., Pollitt, M. (2018): Productivity growth in electricity and gas networks since 1990, Report prepared for the Office of Gas and Electricity Markets (OFGEM), University of Cambridge, Energy Policy Research Group.
- Baischew, D., Cullmann, A., Rechlitz, J., Stronzik, M., Wissner, M. (2022): Ermittlung des generellen Faktorproduktivitätsfortschritts für Gasverteilernetzbetreiber in Österreich im Zuge der vierten Regulierungsperiode, Studie für E-Control Austria.
- Banker, R., Chang, H., 1995. A simulation study of hypothesis tests for differences in efficiencies. *International Journal of Production Economics* 39, 37–54.
- Banker, R.D., H. Chang (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research*, 175(2), 1311-1320
- Behnke, J. (2005): Lassen sich Signifikanztests auf Vollerhebungen anwenden? Einige essayistische Anmerkungen, in: PVS 46, O1–O15.
- Bernstein, J.I., Sappington, D.E.M. (1999): Setting the X Factor in Price-Cap Regulation Plans, in: *Journal of Regulatory Economics* 16(1), 5–26.
- Blum et al. (1999): Grundlagen der Volkswirtschaftslehre, Springer-Verlag.
- Bogetoft, P., Otto, L. (2011): Benchmarking with DEA, SFA, and R, *International Series in Operations Research and Management Science*, Volume 157, Springer-Verlag.
- Braunecker, C. (2016): How to do Empirie, how to do SPSS: Eine Gebrauchsanleitung, Facultas Verlags- und Buchhandels AG.
- Bundesgerichtshof (2021): Beschluss EnVR 7/20 vom 26. Januar 2021 in der energiewirtschaftsrechtlichen Verwaltungssache Genereller sektoraler Produktivitätsfaktor.

- Bundesnetzagentur (2016a): BK4-16-160 vom 05.10.2016.
- Bundesnetzagentur (2016b): BK4-16-161 vom 05.10.2016.
- Bundesnetzagentur (2018a): Beschluss BK4-17-093 vom 21.02.2018.
- Bundesnetzagentur (2018b): Beschluss BK4-18-056 vom 28.11.2018.
- Bundesrat (2007):, Bundesrat Drucksache 417/07, 15.06.07, Verordnung zum Erlass und zur Änderung von Rechtsvorschriften auf dem Gebiet der Energieregulierung, Verordnung der Bundesregierung.
- Bundesverwaltungsgericht (2004): Beschluss vom 21.01.2004 - BVerwG 4 B 82.03.
- Burger, A. (2008) : Produktivität und Effizienz in Banken: Terminologie, Methoden und Status quo, Frankfurt School - Working Paper Series, No. 92, Frankfurt School of Finance & Management.
- CEER [Council of European Energy Regulators] (2020a): CEER Report on Regulatory Frameworks for European Energy Networks 2019, Annex 4, Case studies of single regulatory regimes, Ref: C19-IRB-48-03b, 28 January 2020.
- CEER [Council of European Energy Regulators] (2020b): Report on Regulatory Frameworks for European Energy Networks 2019, Incentive Regulation and Benchmarking Work Stream, Ref: C19-IRB-48-03, 28 January 2020.
- CEPA (2020a): RIIO-GD2 and T2: Cost Assessment – Advice on Frontier Shift policy for Final Determinations, 27 November 2020.
- CEPA (2020b): RIIO-GD2 and T2: Cost Assessment – Frontier shift methodology paper, 27 May 2020.
- Coelli, T., Rao, D.S.P., Battese, G.E. (2002): An introduction to efficiency and productivity analysis, Kluwer Academic Publishers, Sechste Auflage.
- Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G.E. (2005): An introduction to efficiency and productivity analysis. 2nd ed. New York, NY, United States: Springer.
- Conrad, K. (1985): Produktivitätslücken nach Wirtschaftszweigen im internationalen Vergleich, Beschreibung und ökonometrische Ursachenanalyse für die USA, Japan und die Bundesrepublik Deutschland, 1960 – 1979, Springer-Verlag Berlin Heidelberg GmbH.
- Diewert, W.E. (1992): Fisher ideal output, input, and productivity indexes revisited, in: Journal of Productivity Analysis, Volume 3, p. 211–248.
- E-Control (2012): Regulierungssystematik für die zweite Regulierungsperiode GAS 1.1.2013-31.12.2017.
- E-Control (2013): Regulierungssystematik für die dritte Regulierungsperiode der Stromverteilernetzbetreiber 1. Jänner 2014 - 31. Dezember 2018.
- E-Control (2018): Regulierungssystematik für die vierte Regulierungsperiode der Stromverteilernetzbetreiber 1. Jänner 2019 - 31. Dezember 2023.
- E-Control (2018): Regulierungssystematik für die dritte Regulierungsperiode der Gasverteilernetzbetreiber 1. Jänner 2018 - 31. Dezember 2022.
- E-Control (2022): Regulierungssystematik für die vierte Regulierungsperiode der

- Gas-Verteilernetzbetreiber, 1. Jänner 2023 - 31. Dezember 2027.
- Energiavirasto (2021): Sähköverkkoliiketoiminnan kehitys, sähköverkon toimitus varmuus ja valvonnan vaikuttavuus 2020, 4. Mai 2021
- Energimarknadsinspektionen (2022): Reglering av el- och gasnätverksamhet - Utveckling sedan införandet av förhandsregleringen, 25. Januar 2022.
- Franz, O., Stronzik, M. (2005): Benchmarking-Ansätze zum Vergleich der Effizienz von Energieunternehmen, WIK-Diskussionsbeitrag Nr. 262.
- Frontier Economics (2017): Bestimmung des XGEN mittels Törnqvist Index, Ansätze zur Abbildung der Eigenkapitalverzinsung bei der Einstandspreisentwicklung.
- Gugler, K., Liebensteiner, M. (2016): Empirische Schätzung des Produktivitätswachstums und Berechnung des generellen X-Faktors im österreichischen Gasverteilnetz. Wien: Studie für den Fachverband der österreichischen Gas- und Wärmeversorgungsunternehmen, 2016
- Gugler, K., Liebensteiner, M. (2017): „Addendum zu: Empirische Schätzung des Produktivitätswachstums und Berechnung des generellen X-Faktors im österreichischen Gasverteilnetz“. Studie im Auftrag des FGW vom 30. Mai 2017.
- Gugler, K., und M. Liebensteiner (2020): Stellungnahme zur Festlegung des generellen sektoralen Produktivitätsfaktors (GSP) für Betreiber von Stromversorgungsnetzen für die dritte Regulierungsperiode in Deutschland.
- Liebe, A., Schmitt, S., Stronzik, M., Wissner, M. (2017): Gutachten zur Bestimmung des generellen sektoralen Produktivitätsfaktors, Überarbeitete Version nach Eingang der Stellungnahmen, Studie für die Bundesnetzagentur.
- Lohse (1994): Der Rechtsbegriff "Stand der Wissenschaft" aus erkenntnistheoretischer Sicht am Beispiel der Gefahrenabwehr im Immissionsschutz- und Atomrecht, Schriften zum Umweltrecht, Band 36, Duncker & Humblot Berlin.
- Mankiw, N.G., Taylor, P. (2018): Economics, 4th Edition, Cengage Learning.
- McLellan (2004): Measuring Productivity using the Index Number Approach: An Introduction, New Zealand Treasury, Working paper 04/05.
- Meeusen, W., van den Broeck, J. (1977): Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review* 18: 435– 444.
- Meyer, R., Brunekreeft, G., Elias, G. (2020): TOTEX Malmquist index for CPI-X regulation: Does it correctly estimate the true frontier shift?, in: *Journal of Regulatory Economics* (2020) 58:78–97.
- Müller, G. (2009): Produktivitäts- und Effizienzmessung im Eisenbahninfrastruktursektor –Methodische Grundlagen und Schätzung des Produktivitätsfortschritts für den deutschen Markt, WIK-Diskussionsbeitrag Nr. 318, Bad Honnef.
- Naumann, F. (2007): Datenqualität, in: *Informatik Spektrum* 30.1.2007, S. 27 – 31.
- NERA (2020): Frontier Shift at RIIO-T2 Draft Determinations, Prepared for Scottish Power Transmission and National Grid, 2 September 2020.
- Niederprüm, A., Stronzik, M., Wissner, M. (2020): Gutachten zur Ermittlung von Produktivitätsfortschrittsraten, Studie für die Bundesnetzagentur.

- NVE (Norwegian Water Resources and Energy Directorate). (2021): Om den økonomiske reguleringen. abgerufen am 03. Mai 2022 unter <https://www.nve.no/reguleringsmyndigheten/regulering/nettvirksomhet/økonomisk-regulering-av-nettselskap/om-den-okonomiske-reguleringen/>
- OECD (2001): Measuring Productivity, Measurement of aggregate and industry-level productivity growth, OECD Manual.
- Ofgem (2020): RIIO-2 Final Determinations - Core Document, 8. Dezember 2020.
- Oberlandesgericht Düsseldorf (2017): Beschluss vom 20.09.2017 – VI-3 Kart 121/14 (V).
- Oberlandesgericht Düsseldorf (2019): Beschluss vom 10.07.2019 - VI-3 Kart 721/18 (V) .
- Oberlandesgericht Düsseldorf (2022): Beschluss vom 16.03.2022 - VI-3 Kart 637/19 (V).
- Oulton, N.(2016): The Mystery of TFP, in: International Productivity Monitor, No. 31, Fall, S. 68-87.
- Oxera (2017a): Wissenschaftlicher Standard zur Ermittlung des Xgen, Eine Studie für den BDEW vor dem Hintergrund des Festlegungsentwurfs der Bundesnetzagentur (BK4-17-093), 17-November 2017.
- Oxera (2017b): Bestimmung des generellen sektoralen Produktivitätsfortschritts für Strom- und Gasnetzbetreiber, Untersuchung für den BDEW Bundesverband der Energie- und Wasserwirtschaft e.V., 6. Februar 2017.
- Parman, B.J., Featherstone, A.M. (2019): A comparison of parametric and nonparametric estimation methods for cost frontiers and economic measures, in: Journal of Applied Economics, 2019, VOL. 22, No. 1, 60–85.
- Polynomix und Jacobs University (2016): Die Ermittlung des technologischen Fortschritts anhand von Unternehmensdaten. Der Einsatz der Malmquist-Methode im deutschen Regulierungsrahmen. Studie im Auftrag der Netze BW GmbH, August 2016.
- Sachs,L., Hedderich, J. (2009): Angewandte Statistik, Methodensammlung mit R, Dreizehnte aktualisierte und erweiterte Auflage, Springer-Verlag.
- Schmitt, S., Stronzik, M. (2015): Die Rolle des generellen X-Faktors in verschiedenen Regulierungsregimen, WIK-Diskussionsbeitrag Nr. 399.
- Schmitt, S., Stronzik, M., Wissner, M. (2017): Ermittlung des generellen Faktorproduktivitätsfortschritts für Gasverteilernetzbetreiber in Österreich im Zuge der dritten Regulierungsperiode, Studie der WIK-Consult für die E-Control.
- Schweinsberg, A., Stronzik, M., Wissner, M. (2011): Cost Benchmarking in Energy Regulation in European Countries – Study of WIK-Consult for the Australian Energy Regulator.
- Statistisches Bundesamt (2008): Klassifikation der Wirtschaftszweige, mit Erläuterungen.
- Statistisches Bundesamt (2017): Verdienste und Arbeitskosten, Index der Tarifverdienste und Arbeitszeiten, 4. Vierteljahr 2016.
- Statistisches Bundesamt (2018): Preise, Verbraucherpreisindex für Deutschland, Qualitätsbericht, erschienen am 04.01.2018.
- Statistisches Bundesamt (2019): Handbuch zur Methodik Index der Erzeugerpreise gewerblicher Produkte (Inlandsabsatz).
- Statistisches Bundesamt (2021): Arbeitskostenindex, 2020, erschienen am 12.05.2021.

- Statistisches Bundesamt (2022): Index der Erzeugerpreise gewerblicher Produkte (Inlandsabsatz) nach dem Güterverzeichnis für Produktionsstatistiken, Ausgabe 2009 (GP 2009), Lange Reihen der Fachserie 17, Reihe 2 von Januar 2005 bis Februar 2022.
- Stronzik, M., Wissner, M. (2018): Ermittlung des generellen Faktorproduktivitätsfortschritts für Stromverteilernetzbetreiber in Österreich im Zuge der vierten Regulierungsperiode, Studie für E-Control Austria.
- Süchting, J. (1995): Finanzmanagement, Theorie und Politik der Unternehmensfinanzierung, Springer-Verlag.
- Trinkner, U., Mattmann, M. (2018): Törnqvist-Tool Xgen RP3 Strom, Prüfbericht im Auftrag der Bundesnetzagentur.
- Von Auer, L. (2019): Gutachten nach Beweisbeschluss des Oberlandesgerichts vom 17.09.2018 und Schreiben vom 17.09.2018. Sachverständigengutachten vom 4.2.2019.
- Wangsness, P.B., Halse, A.H. (2021): The Impact of Electric Vehicle Density on Local Grid Costs: Empirical Evidence from Norway. The Energy Journal, Vol. 42, No. 5.

8 Anhang: Berechnungen für den Gassektor anhand des Excel-Törnqvist-Tools

Für die Festlegung des sektoralen Produktivitätsfortschritts für die vierte Regulierungsperiode verwendet die Bundesnetzagentur ein Excel-Tool. Im Folgenden wird es auf Konsistenz und Umsetzung überprüft und plausibilisiert. Die Überprüfung und Plausibilisierung enthält sämtliche Tabellenblätter, inkl. aller vorhandenen Formeln und Bezüge. Die Plausibilisierungen beziehen sich in Abschnitt 8.2 stets auf den Zeitraum 2006 bis 2021. Dabei benennt das Jahr immer die Veränderungsrate von diesem Jahr zum Vorjahr, das Jahr 2011 beispielsweise benennt also die Veränderung zum Jahr 2010. Aus diesem Grund ist eine Berechnung mit dem Startjahr 2006 nicht möglich (da keine Werte für 2005 hinterlegt sind). Die Veränderungsrate 2007 erfasst aber die Veränderung von 2006 nach 2007.

Im Folgenden werden die einzelnen Tabellenblätter erläutert und einer Prüfung unterzogen.

8.1 Datengrundlage

Die Summen der Datenblätter „00 GuV“ und „00 SAV“ wurden überprüft. Es wurden die Daten aus den Datenerhebungsbögen der Netzbetreiber für die Tabellenblätter „Datenabfrage“, „Netzbetreiber_Anlagevermögen“ und „Verpächter_Anlagevermögen“ in Excel aggregiert und mit den vorläufigen Berechnungen der Bundesnetzagentur verglichen. So konnten potenzielle Abweichungen festgestellt werden. Auf diese Weise konnten unplausible Werte eliminiert werden, indem offensichtlich falsche Datensätze korrigiert wurden und eine Neuberechnung stattfand.

8.2 Technische Umsetzung

8.2.1 Tabellenblatt „01_Xgen“

Auf diesem Tabellenblatt wird der generelle sektorale Produktivitätsfaktor in den Regulierungskontext eingebettet. Die für die Berechnung herangezogene Formel wird erläutert. Dabei wird die sog. Residualbetrachtung angewendet. Eine ausführliche Auseinandersetzung darüber findet sich in Abschnitt 3.

Der Wert für den Xgen wird als Ergebnis der Berechnungen aus den anderen Tabellenblättern korrekt dargestellt. Die Bezüge aus den anderen Tabellenblättern erfolgen korrekt.

Für den Xgen ergibt sich auf diese Weise ein Wert von **0,75 %**.

8.2.2 Tabellenblatt „01_Inputpreisentwicklung_Netz“

Auf diesem Tabellenblatt erfolgt die Berechnung der Inputpreise der Gasnetzbetreiber. Die verwendeten Zeitreihen basieren auf Zeitreihen von Destatis, der Deutschen Bundesbank und aus der Datenabfrage der Bundesnetzagentur an die Netzbetreiber.

Die Start- und Endjahreingabe erfolgt in den Zellen B1 und B2: Die beiden Zellen werden in den Zellen B27 bzw. T23 über das Tabellenblatt „Hilfstabelle“ richtig zur Berechnung des VPI und IP Energie zugeordnet. Dabei benennt – wie oben beschrieben - das Jahr immer die Veränderungsrate von diesem Jahr zum Vorjahr.

Plausibilisierung:

Die Berechnung von TF_{Netz} , IP_{Energie} und VPI und erfolgt jeweils als geometrisches Mittel über den gesamten Zeitraum. Alternativ könnten für den Xgen auch zunächst die einzelnen Jahre berechnet und anschließend das geometrische Mittel über die jährlichen Xgen-Werte gebildet werden. Das Ergebnis für den Xgen beträgt dann 0,71 %.

Eine solche Berechnung wäre allerdings nicht überlegen, da für den generellen sektoralen Produktivitätsfaktor die Ermittlung eines aggregierten Indexes über die Zeit im Vordergrund steht und weniger die Werte der Einzeljahre. Dies ergibt sich bereits aus § 9 Abs. 1 ARegV, nach dem die Abweichungen des Produktivitätsfortschritts und der Einstandspreisentwicklungen gemessen werden soll, nicht die Entwicklung der Abweichungen der Einzeljahre.

Die Berechnung des gesamten Xgen (Zelle E1) wird logisch korrekt vom Tabellenblatt „01_Xgen“ bezogen. Die Indexreihen der einzelnen Zeitreihen sind logisch korrekt verlinkt.

Die Berechnungen der einzelnen Gesamtveränderungsfaktoren pro Jahr erfolgt in den Zellen T22 bis AH22. Das Produkt der Einzelveränderungsfaktoren ergibt den Gesamtwert (nach Törnqvist). Die Berechnung der Einzelveränderungsfaktoren pro Jahr (Zellen T5 bis AH19) erfolgt mathematisch korrekt nach Törnqvist mit korrekt gewichteten jährlichen Veränderungsrate (Gewichtung gemäß Anteilen an Gesamtkosten in Arbeitsblatt „02_GuV_Anteile“ korrekt verlinkt).

8.2.3 Tabellenblatt „01_Produktivitätsentw._Netz“

Auf diesem Tabellenblatt erfolgt die Berechnung der Totalen Faktorproduktivität bzw. des technischen Fortschritts des Netzsektors.

Die Start- und Endjahre in den Zellen B1 und B2 werden korrekt aus der Tabelle „01_Inputpreisentwicklung_Netz“ übernommen.

In Zelle D48 erfolgt die Berechnung der Veränderung der Produktivität korrekt als geometrisches Mittel des gewählten Betrachtungszeitraums (Startjahr / Endjahr) über die Hilfstabelle.

Plausibilisierung:

Die Berechnung der TFP erfolgt jeweils als geometrisches Mittel der einzelnen Jahreswerte (Outputindex dividiert durch Inputindex) über den Betrachtungszeitraum; alternativ könnten zunächst das geometrische Mittel von Input- und Outputindex über den gesamten Zeitraum berechnet und anschließend der Quotient der beiden Werte gebildet werden. Durch Anwendung des geometrischen Mittels besteht dadurch kein Unterschied im Ergebnis (vgl. auch Abschnitt 2.3.3).

Die jahresweise Berechnung des TF findet in den Zellen D44 bis R44 statt, die Division des Outputindex durch den Inputindex erfolgt korrekt.

Die jahresweise Berechnung des **Outputindexes** erfolgt in den Zellen D28 bis R28: Die Berechnung der Indexreihe erfolgt mit Bezug auf Zelle A26 (Auswahl des Outputparameters), falls der Output dem Bruttoproduktionswert (preisbereinigt) oder der Bruttowertschöpfung (preisbereinigt) entspricht.

In Zelle A26 sind also zwei verschiedene Einstellungen möglich:

- Bruttoproduktionswert (preisbereinigt):
 - Über Verweis-Funktion wird die Zeitreihe (nämlich C9 bis N9) korrekt gefunden und in die Zellen C26 bis R26 übernommen.
 - Der Bruttoproduktionswert (preisbereinigt) selbst wird in den Zellen C9 bis R9 als Summe aus Umsatzerlösen (preisbereinigt) (Zellen C11 bis R11), den Bestandsveränderungen (preisbereinigt) (Zellen C13 bis R13) und den aktivierten Eigenleistungen (preisbereinigt) (Zellen C15 bis R15) gebildet.
 - Eine Berechnung eines Törnqvist-Indexes ist aufgrund teilweise negativer Werte in der Reihe „Bestandsveränderungen“ nicht möglich (negative Anteile). Daher erfolgt die Summenbildung.
 - Die genannten Summanden werden logisch korrekt aus dem Tabellenblatt „02_Deflatoren“ übernommen.
- Bruttowertschöpfung (preisbereinigt):
 - Über Verweis-Funktion wird die Zeitreihe (nämlich C24 bis R24) korrekt gefunden und in die Zellen C26 bis R26 übernommen.
 - Die Bruttowertschöpfung selbst wird in den Zellen C24 bis R24 sachgerecht als Differenz aus Bruttoproduktionswert (preisbereinigt) (Zellen C9 bis R9) und Aufwendungen für RHB (preisbereinigt) (Zellen C18 bis R18), Aufwendungen für bezogene Leistungen (preisbereinigt) (Zellen C20 bis R20) sowie den sonstigen betrieblichen Aufwendungen (preisbereinigt) (Zellen C21 bis R21) gebildet.
 - Die genannten Subtrahenden werden logisch korrekt aus dem Tabellenblatt „02_Deflatoren“ übernommen.

Plausibilisierung:

Eine Berechnung mit der Bruttowertschöpfung als Outputgröße ergibt für den Xgen einen Wert von 0,76 %. Dieser Wert liegt knapp über dem Wert des Xgen mittels einer

Berechnung mit dem Produktionswert als Outputgröße und entspricht somit den Erwartungen aus der wissenschaftlichen Literatur. Weitere Erläuterungen zu diesem Aspekt finden sich in 4.2.2.

Die jahresweise Berechnung des **Inputindex** findet in den Zellen D42 bis R42 statt:

Zunächst erfolgt eine Abfrage, ob die Bruttowertschöpfung oder der Bruttoproduktionswert betrachtet wird. Die Berechnung der Veränderungsrate erfolgt sachgerecht nach der Törnqvist-Formel mit der korrekten Gewichtung der Inputfaktoren. Die Inputs ergeben sich aus den Reihen Bruttoanlagevermögen (TNW 2015) (C30 bis R30), geleistete Arbeitsstunden (C32 bis R32) sowie den Vorleistungen (Preisbereinigt) (C34 bis R34). Die Anteile der Inputfaktoren ergeben sich aus den Reihen C36 bis R36 (Lohnquote), C40 bis R40 (Kapitalquote) sowie C38 bis R38 (Vorleistungsquote). Letztere fällt im Falle der Ansetzung der BWS korrekterweise auf 0, wobei sich Lohn- und Kapitalquote auf 100 % summieren.

Die genannten eingehenden Berechnungsgrößen Bruttoanlagevermögen (TNW 2015) (C30 bis R30) und geleistete Arbeitsstunden (C32 bis R32) werden logisch korrekt aus den anderen Tabellenblättern bezogen; die Vorleistungen (Preisbereinigt) (C34 bis R34) sachlich richtig im Tabellenblatt berechnet.

Die Berechnung der **Lohnquote** erfolgt in den Zellen C36 bis R39.

Zunächst erfolgt eine korrekte Abfrage, ob die Bruttowertschöpfung oder der Bruttoproduktionswert betrachtet wird.

Im Falle des Bruttoproduktionswerts erfolgt die Berechnung durch Division des logisch korrekt aus dem Tabellenblatt „02_Personalaufwand“ bezogenen Personalaufwands durch die Summe der korrekterweise nicht preisbereinigten Größen „Umsatzerlöse“ (Zellen C10 bis R10), „Bestandsveränderungen“ (Zellen C12 bis R12) und „Aktivierte Eigenleistungen“ (Zellen C14 bis R17).

Im Falle der BWS erfolgt die Berechnung durch Division des logisch korrekt aus dem Tabellenblatt „02_Personalaufwand“ bezogenen Personalaufwands durch die Summe der korrekterweise nicht preisbereinigten Größen „Umsatzerlöse“ (Zellen C10 bis R10), „Bestandsveränderungen“ (Zellen C12 bis R12) und „Aktivierte Eigenleistungen“ (Zellen C14 bis R14) abzüglich der Größen:

- Aufwendungen für RHB (korrekterweise nicht preisbereinigt) (Zellen C17 bis R17)
- Aufwendungen für bezogene Leistungen (korrekterweise nicht preisbereinigt) (Zellen C19 bis R19) und
- Sonstige betriebliche Aufwendungen logisch korrekt bezogen aus Tabellenblatt „02_Deflatoren“ (Zellen C36 bis R36).
- Die sonstigen betriebliche Aufwendungen (nicht preisbereinigt) könnten aus Gründen der Konsistenz ebenfalls in das Tabellenblatt „01_Produktivitätswertw. Netz“ aufgenommen werden. Da dies allerdings nicht ergebnisrelevant ist,

wird

darauf

verzichtet.

Die Berechnung der **Vorleistungsquote** erfolgt in den Zellen C38 bis N38:

Zunächst erfolgt eine korrekte Abfrage, ob die Bruttowertschöpfung oder der Bruttoproduktionswert betrachtet wird.

Im Falle des Bruttoproduktionswerts erfolgt die Berechnung richtigerweise durch Division der Summe der Größen „Aufwendungen für RHB“ (Zellen C17 bis R17; korrekterweise nicht preisbereinigt), „Aufwendungen für bezogene Leistungen“ (Zellen C19 bis R19; korrekterweise nicht preisbereinigt) und der Größe „Sonstige betriebliche Aufwendungen“ (logisch korrekt bezogen aus dem Arbeitsblatt „02_Deflatoren, Zellen C36 bis R36; korrekterweise nicht preisbereinigt) durch die Summe der korrekterweise nicht preisbereinigten Größen „Umsatzerlöse“ (Zellen C10 bis R10), „Bestandsveränderungen“ (Zellen C12 bis R12) und „Aktivierte Eigenleistungen“ (Zellen C14 bis R14).

Im Falle der BWS erfolgt die Ansetzung der Vorleistungsquote korrekterweise zu einem Wert von 0.

Die Berechnung der **Kapitalquote** erfolgt in den Zellen C40 bis R40.

Zunächst erfolgt eine korrekte Abfrage, ob die Bruttowertschöpfung oder der Bruttoproduktionswert betrachtet wird. Im Falle des Bruttoproduktionswerts erfolgt die Berechnung richtigerweise als Residualgröße (1 abzgl. Lohnquote abzgl. Vorleistungsquote). Im Falle der BWS erfolgt die Berechnung richtigerweise als Residualgröße (1 abzgl. Lohnquote).

8.2.4 Tabellenblatt „02_Deflatoren“

Die Abfrage der Umsatzerlöse (Zellen C5 bis R5) erfolgt logisch und sachlich korrekt aus dem Datenblatt „00_GuV“.

Die Preisbereinigung der **Umsatzerlöse** erfolgt in den Zellen C6 bis R6. Als Deflator werden gewichtete Netzentgelte herangezogen. Zunächst erfolgt eine Gewichtung der verschiedenen Netzentgelte (Netzentgelte Strom inkl. Messung, Abrechnung, MSB in ct/kWh [Haushaltskunde 3.500 kWh]; Netzentgelte Strom inkl. Messung, Abrechnung, MSB in ct/kWh [Gewerbekunde 50 MWh]; Netzentgelte Strom inkl. Messung, Abrechnung, MSB in ct/kWh [Industriekunde 24 GWh]) anhand des Stromverbrauchs dieser Verbrauchsgruppen. Entsprechend ihrer Gewichtung (Zellen B9 bis B11) gehen die verschiedenen Netzentgelte in ein Gesamtnetzentgelt (Zellen C7 bis R7) ein. Die Umsatzerlöse werden mathematisch korrekt preisbereinigt.

Die Abfrage der **Bestandsveränderungen** (Zellen C14 bis R14) erfolgt logisch und sachlich korrekt aus dem Datenblatt „00_GuV“. Die Preisbereinigung der Bestandsveränderungen erfolgt in den Zellen C15 bis R15: Als Deflator für Bestandsveränderungen wird logisch korrekt auf die Indexwerte in den Zellen C25 bis R25 verwiesen. Die Bestandsveränderungen werden mathematisch korrekt preisbereinigt.

Die Abfrage der **aktivierten Eigenleistungen** (Zellen C18 bis R18) erfolgt logisch und sachlich korrekt aus dem Datenblatt „00_GuV“. Die Preisbereinigung der aktivierten Eigenleistungen erfolgt in den Zellen C19 bis R19:

Als Deflator für aktivierte Eigenleistungen wird logisch korrekt auf die Indexwerte in den Zellen C20 bis R20 verwiesen. Der verwendete Index (Personalaufwand / tatsächlich geleistete Arbeitsstunden) wird korrekt aus dem Tabellenblatt „Inputpreisentwicklung_Netz“ übertragen. Die aktivierten Eigenleistungen werden mathematisch korrekt preisbereinigt.

Die Abfrage der **RHB** (Zellen C22 bis R22) erfolgt logisch und sachlich korrekt aus dem Datenblatt „00_GuV“. Die Preisbereinigung der RHB erfolgt in den Zellen C23 bis R23:

Es wird ein gewichteter Gesamtindex aus den Indexreihen „Index der Erzeugerpreise gewerblicher Produkte (Vorleistungsgüterproduzenten und Investitionsgüterproduzenten)“ und „Elektrischer Strom an Weiterverteiler“ gebildet. Die Indexreihen werden logisch korrekt bezogen. Der erstgenannte Index wird je zu 50 % aus den Indexreihen „Index der Erzeugerpreise gewerblicher Produkte; Erzeugnisse der Vorleistungsgüterproduzenten“ und „Index der Erzeugerpreise gewerblicher Produkte (Inlandsabsatz); Erzeugnisse der Investitionsgüterproduzenten“ gebildet. Dies ist vor dem Hintergrund zu sehen, dass es sich bei den RHB zum Teil um aktivierungsfähige Aufwendungen für das Sachanlagevermögen handelt, die aber als Aufwand gebucht wurden. Daher ist es stringent, auch den Index für Investitionsgüter einzubeziehen. Der Ansatz des Indexes der Erzeugerpreise gewerblicher Produkte (Vorleistungsgüterproduzenten und Investitionsgüterproduzenten) für diese Positionen ist sachgerecht, da dieser Index zur Abbildung der Preisentwicklung der übrigen Roh-, Hilfs- und Betriebsstoffe geeignet ist. Die RHB werden mathematisch korrekt preisbereinigt.

Die Abfrage der Aufwendungen für **bezogene Leistungen** (Zellen C29 bis R29) erfolgt logisch und sachlich korrekt aus dem Datenblatt „00_GuV“. Die Preisbereinigung der Aufwendungen für bezogene Leistungen erfolgt in den Zellen C30 bis R30: Es wird ein gewichteter Gesamtindex aus den Indexreihen „Arbeitskostenindex Produzierendes Gewerbe und Dienstleistungsbereich“ und „Index der Erzeugerpreise Erdgas, bei Abgabe an Wiederverkäufer“ gebildet. Die Indexreihen werden logisch und inhaltlich korrekt bezogen.

Die Bildung des Gesamtindexes ist korrekt, der Arbeitskostenindex Produzierendes Gewerbe und Dienstleistungsbereich wird auf das Jahr 2015 normiert, um ihn mit dem Erzeugerpreise Erdgas, bei Abgabe an Wiederverkäufer, zu kombinieren.

- Die Abfrage der **sonstigen betrieblichen Aufwendungen** (Zellen C36 bis R36) erfolgt logisch und sachlich korrekt aus dem Datenblatt „00_GuV“.

Die Preisbereinigung der Aufwendungen für bezogene Leistungen findet in den Zellen C37 bis R37 statt. Die Indexreihe (Zellen C38 bis R38) wird logisch korrekt bezogen.

Die Aufwendungen für bezogene Leistungen werden mathematisch korrekt preisbereinigt. (Bezugsjahr 2015).

8.2.5 Tabellenblatt „02_Index_Betr_Aufwendungen“

Es wird ein Mischindex für verschiedene Dienstleistungen gebildet. Die Bildung des Index erfolgt über die Gewichtung von Einzelindizes. Der Bezug der Einzelindizes aus anderen Tabellenblättern ist logisch und inhaltlich korrekt.

8.2.6 Tabellenblatt „02_GuV_Anteile“

Diese Tabelle bildet die Grundlage für die Gewichtungen der Inputpreisbestandteile. Die Bezüge der einzelnen Größen und Berechnungen sind logisch, mathematisch und inhaltlich korrekt.

8.2.7 Tabellenblatt „02_Anlagevermögen“

Es erfolgt die Berechnung der Tagesneuwerte in Preisen von 2015 für die verschiedenen Jahre (2006 bis 2021). Dazu werden die Vermögensgegenstände über das Tabellenblatt „Faktorreihen“ korrekt mit den Indexreihen verknüpft. Für die Anlagengruppe 1 (Grundstücke) findet nach § 6 GasNEV i.V.m. Anlage 1 zur GasNEV keine Tagesneuwertberechnung statt.

Die Anlagen sind aufgrund der durchgeführten Preisbereinigung (TNW 2015) auf das gleiche Basisjahr bezogen. Die Verlinkung auf das Tabellenblatt „00_SAV“ erfolgt mathematisch und logisch korrekt.

8.2.8 Tabellenblatt „02_Personalaufwand“

Es erfolgt die Berechnung des

- Personalaufwands
- Personalaufwands (preisbereinigt)
- Personalaufwands / tatsächlich geleistete Arbeitsstunden

Die Berechnungen / Bezüge sind logisch und inhaltlich korrekt.

8.2.9 Hilfstabelle

Die Hilfstabelle ist inhaltlich korrekt.

8.2.10 Folgende Tabellenblätter erhalten Daten aus öffentlich zugänglichen Quellen:

- 02_VPI
- 02_Erzeugerpreise:

Die Indexreihen sind korrekt aus der Statistik übernommen (Lange Reihen der Fachserie 17, Reihe 2 von Januar 2005 bis März 2022, S. 4).

- 02_Erzeugerpreise_2
- 02_Arbeitskostenindizes
- 02_Zinsreihen
- 02_Gewerbesteuerhebesätze
- 02_BWS
- 02_BAV
- 02_Arbeitsstunden
- 02_AN-Entgelte
- 02_Vorleistungen
- 02_Produktionswert
- 02_Dienstleistungen
- 02_VPI_Detail
- 02_Großhandelsverkaufspreise

8.2.11 Folgende Tabellenblätter erhalten jährliche Mittelwertbildungen.

Diese sind mathematisch korrekt:

- BBK01.WU0004_Anleihen_öH
- BBK01.WU0022_Anleihen_Untern
- BBK01.WU0018_Hypotheken
- BBK01.WU001_Umlaufrendite

Folgende Tabellenblätter dienen rein als Datengrundlage und wurden nicht inhaltlich geprüft:

- Anlagengruppen
- Faktorreihen
- Zuordnung_Faktorreihen

8.3 Weitere Plausibilisierungen

Unter der Beachtung eines Mindestzeitraums von vier Jahren, entsprechend der Vorgabe des § 9 Abs. 3 S. 2 ARegV, lassen sich mit dem Excel-Tool Xgen-Werte für verschiedene Zeitintervalle berechnen. Nachstehend wurde daher mit verschiedenen Ansätzen untersucht, ob das auf Grundlage des Törnquistindex für die vierte Regulierungsperiode im Gasbereich bislang ermittelte Ergebnis für den Gasbereich plausibel erscheint. In den nachfolgenden Ausführungen benennt das Jahr immer die Veränderungsrate von diesem Jahr zum Vorjahr. Das Jahr 2011 benennt also die Veränderung vom Jahr 2010 zum Jahr 2011

Insoweit kommt zunächst folgende Betrachtungsweise in Betracht: Unter den Ergebnissen findet sich zwangsläufig der bisher auf Grundlage des Törnqvist ermittelte Wert in Höhe von 0,75 % als Ergebnis des längstmöglichen Stützintervalls (vgl. hierzu Abschnitt 2.3.2). Bei dieser Betrachtung fungiert die Entwicklung vom Jahr 2006 auf das Jahr 2007 als Startjahr:

Tabelle 8-1: Xgen-Werte für verschiedene Zeiträume I (in %)

Zeitraum	2007 -2009	2007 -2010	2007 -2011	2007 -2012	2007 -2013	2007 -2014	2007 -2015
Xgen	-3,31	-1,07	-2,12	2,23	-0,05	-0,21	-0,53

Zeitraum	2007 -2016	2007 -2017	2007 -2018	2007 -2019	2007 -2020	2007 -2021
Xgen	0,49	0,32	0,01	0	-0,25	0,75

Es ergeben sich (mit Ausnahme des Wertes für den Zeitraum 2007 bis 2012) relativ geringe positive und negative Werte für den Xgen. Wie zu erkennen ist, steigt der Xgen tendenziell, je näher die einhegenden Werte an die Gegenwart und hin zum längstmöglichen Stützintervall rücken.

Darüber hinaus besteht die Möglichkeit, den näher an der Gegenwart liegenden Zeiträumen ein stärkeres Gewicht eingeräumt wird. Das Endjahr 2021 bleibt dann fixiert, während die Anfangsjahre im Zeitablauf steigen. Auch hier ist der Zeitraum 2007 bis 2021 enthalten. Es ergibt sich auf diese Weise folgendes Bild.

Tabelle 8-2: Xgen-Werte für verschiedene Zeiträume II (in %)

Zeitraum	2007 -2021	2008 -2021	2009 -2021	2010 -2021	2011 -2021	2012 -2021	2013 -2021
Xgen	0,75	-0,98	0,34	1,78	1,41	2,2	-0,22

Zeit- raum	2014 -2021	2015 -2021	2016 -2021	2017 -2021	2018 -2021	2019 -2021
Xgen	1,44	1,83	2,66	1,28	1,92	3,72

In diesem Fall liegt der Wert von 0,75 % unter dem Durchschnitt aller Werte (1,39 %) für die verschiedenen Zeiträume. Würde man den Xgen also stärker auf kürzer in der Vergangenheit liegende Jahre basieren, so läge dieser im Durchschnitt höher.

Schließlich kann – wie in Abschnitt 2.3.4 diskutiert – eine weitere Methode zur Validierung in der Fortschreibung des Trends aus den Daten der Vergangenheit liegen. Würde man den Trend für den Xgen linear fortschreiben, so ergäben sich folgende Werte (vgl. Tabelle 8-3).

Tabelle 8-3: Xgen-Werte für zukünftige Jahre (in %)

Jahr	2022	2023	2024	2025	2026	2027
Xgen	2,24 %	2,33 %	2,42 %	2,51 %	2,60 %	2,69 %

Diese Werte liegen deutlich und einheitlich über dem bislang ermittelten Wert, der sich unter Heranziehung des längstmöglichen Stützintervalls ergibt.