

Stocker | Birke | Brunekreeft | Weiß [Hrsg.]

Perspektiven der Netzökonomie

Beiträge zu Ehren von Günter Knieps



Nomos

Freiburger Studien zur Netzökonomie

Herausgegeben von Prof. Dr. Germain Gaudin

Diese Schriftenreihe wurde begründet von
Prof. Dr. Günter Knieps

Band 23 | Volume 23

Volker Stocker | Franziska Birke

Gert Brunekreeft | Hans-Jörg Weiß [Hrsg.]

Perspektiven der Netzökonomie

Beiträge zu Ehren von Günter Knieps



Nomos

Gedruckt mit freundlicher Unterstützung der Gesellschaft für Netzökonomie, Wettbewerbsökonomie und Verkehrswissenschaft an der Universität Freiburg.

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>

ISBN 978-3-7560-0497-3 (Print)
978-3-7489-3746-3 (ePDF)

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN 978-3-7560-0497-3 (Print)
978-3-7489-3746-3 (ePDF)

Library of Congress Cataloging-in-Publication Data

Stocker, Volker | Birke, Franziska | Brunekreeft, Gert | Weiß, Hans-Jörg

Perspektiven der Netzökonomie

Beiträge zu Ehren von Günter Knieps

Volker Stocker | Franziska Birke | Gert Brunekreeft | Hans-Jörg Weiß (Eds.)

231 pp.

Includes bibliographic references.

ISBN 978-3-7560-0497-3 (Print)
978-3-7489-3746-3 (ePDF)

1. Auflage 2023

© Nomos Verlagsgesellschaft, Baden-Baden 2023. Gesamtverantwortung für Druck und Herstellung bei der Nomos Verlagsgesellschaft mbH & Co. KG. Alle Rechte, auch die des Nachdrucks von Auszügen, der fotomechanischen Wiedergabe und der Übersetzung, vorbehalten. Gedruckt auf alterungsbeständigem Papier.

This work is subject to copyright. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without prior permission in writing from the publishers. Under § 54 of the German Copyright Law where copies are made for other than private use a fee is payable to "Verwertungsgesellschaft Wort", Munich.

No responsibility for loss caused to any individual or organization acting on or refraining from action as a result of the material in this publication can be accepted by Nomos or the editors.

Geleitwort

Carl Christian von Weizsäcker

Günter Knieps kam auf dem Umweg über Werner Hildenbrand von der Mathematik zur angewandten Ökonomie. Als er nach seiner Promotion an der Universität Bonn in den späten siebziger Jahren die „reine Theorie“ verließ, wurde er unter meiner Leitung Mitglied eines kleinen Teams von Volkswirten, die sich mit der Thematik der leistungsgebundenen Wirtschaftsleistungen befasste. Dieses Thema wurde damals aktuell wegen eines sich anbahnenden technologischen Umbruchs, den wir heute die Digitalisierung nennen. Zu den betroffenen Branchen gehörten insbesondere die Telekommunikation und die Stromversorgung, ferner die sich im Aufbau befindende Erdgasversorgung.

Die Fragestellungen, die wir aus wirtschaftspolitischer Sicht behandelten, waren recht neu. Und so konnte man mit nur geringen Branchenvorkenntnissen recht schnell in Forschungsneuland vorstoßen. Im Rahmen dieser Mannschaft gab es eine Weile lang ein Unter-Team bestehend aus Günter Knieps, Jürgen Müller und mir. Wir bekamen seitens der Monopolkommission den Auftrag, ein Gutachten über die Möglichkeit zu erstellen, die dem Wettbewerb in der Telekommunikation geboten werden konnten. Günter arbeitete sich schnell insbesondere in die USA-Literatur ein. Dann zogen wir zu dritt los, um in Interviews den Stand der Dinge abzuklären, so bei den Lieferanten der Telekom, bei den großen Kunden und anderen an diesen Fragen Interessierten. Wir lieferten unser Gutachten an die Monopolkommission pünktlich ab. Dort stieß es auf Resonanz; aber auch in der Öffentlichkeit. Unsere Arbeitsgruppe an der Universität Bonn ging weiter. Wir wurden auch Berater des damaligen Bundespostministers Matthöfer, dem wir helfen sollten, einen hausinternen „Think Tank“ aufzubauen. Aus dem ist später das WIK geworden, das heute noch als Forschungs- und Beratungsinstitut sehr erfolgreich fungiert.

Ich entschloss mich im Jahre 1981 an die Universität Bern überzusiedeln. Eine kleine Mannschaft von Mitarbeitern nahm ich mit, so auch Günter Knieps. Karl Heinz Neumann blieb in Bonn und wurde der Gründungsdirektor des Think Tanks im Bundespostministerium. Günter Knieps blieb als mein Assistent dem Thema Telekommunikation treu. Wir haben gemeinsam die regelmäßige Vorlesung „Verkehrspolitik und

Telekommunikation“ angeboten, die bei den Studenten und Studentinnen auch viel Anklang fand. Dies bot Günter Knieps auch die Gelegenheit, sich in die Verkehrsverhältnisse des Alpenlandes Schweiz einzuarbeiten. Aus seinem Hauptarbeitsgebiet der Telekommunikation entstand seine Habilitationsschrift, die eine exzellente Analyse der Wettbewerbsregulierung des US-amerikanischen Telekommunikationswesens darstellt. Es war dies die Zeit des Antitrust-Prozesses gegen das Telekommunikationsmonopol ATT – und damit eine Zeit großer Kämpfe um die beste Telekommunikationspolitik. Ich selbst wurde von der schweizerischen Bundesregierung als Berater und Gutachter in Sachen Telekommunikation aufgeboten – und konnte dank der kompetenten Unterstützung durch Günter Knieps vielleicht ganz gute Beratung geben. Es waren viele veraltete Auffassungen durch geduldige Überzeugungsarbeit zu überwinden. Besonders interessant war das Suchen nach Kompromissen zwischen dem an digitalen Reformen höchst interessierten Finanzsektor und der dem Hergebrachten anhängenden Bevölkerung des ländlichen Raums.

Nach seiner Habilitation an der Universität Bern war Knieps recht bald ein gefragter Experte. Ein Ruf nach Groningen ermöglichte ihm, auch die in vielerlei Hinsicht sehr speziellen Niederlande kennen zu lernen und dort auch als Regierungsberater tätig zu werden. Sodann kam der Ruf an die Universität Freiburg im Breisgau – an den Fachbereich, der nicht zuletzt auch die Tradition des Ordo-Liberalismus weiterführte. Hier konnte Günter ein gutes Team aufbauen, das in den Bereichen Telekommunikation, Verkehr und Energie arbeitete. Inzwischen war das allgemeine Interesse an der digitalisierungsgetriebenen Telekommunikation auch in Deutschland sehr groß geworden. Die ausgezeichneten Arbeiten von Günter Knieps und seines Teams auf diesem Gebiet machten Furore. So wurde Günter auch sowohl in den wissenschaftlichen Beirat des Bundesverkehrsministeriums als auch in den des Bundeswirtschaftsministeriums berufen. Im Rahmen dieses letzteren Beirats ergab sich für mich auch eine erneute Kooperation mit Günter Knieps.

Lieber Günter – herzlichen Dank für die vielen Jahre freundschaftlicher wissenschaftlicher Zusammenarbeit.

Inhalt

Einführungsbeitrag der Herausgeber <i>Franziska Birke, Gert Brunekreeft, Volker Stocker und Hans-Jörg Weiß</i>	9
Agile Regulation: Experiments in Flexible Governance <i>Johannes M. Bauer</i>	31
Preisbildung bei der Zusammenschaltung im Internet <i>Falk von Bornstaedt</i>	51
Networks, Centrality, Public Good Index and the Medici <i>Manfred J. Holler & Florian Rupp</i>	73
Lessons learned – Anstöße der Regulierung zum Abschreibungsproblem <i>Hans-Ulrich Küpper</i>	89
Der Deutschland-Takt: Regulierung, Wettbewerb und Organisation auf der Schiene <i>Kay Mitusch</i>	107
Der Beitrag von (nicht-)ökonomischen Faktoren zur Analyse von Korruption, Steuerhinterziehung, Steuermoral und Schwarzarbeit <i>Friedrich Schneider</i>	135
Wettbewerbspolitik und Demokratie: Das Spannungsfeld von Konsumenten-, Produzenten- und Bürgerinteressen <i>Viktor J. Vanberg</i>	159
Die Anti-Stau-Maut <i>Achim Wambach</i>	177

Inhalt

Der disaggregierte Regulierungsansatz und die preußischen Eisenbahnen im 19. Jahrhundert	195
<i>Bernhard Wieland</i>	
Access regulation of non-personal data markets	211
<i>Patrick Zenhäusern</i>	

Einführungsbeitrag der Herausgeber

Franziska Birke¹, Gert Brunekreeft², Volker Stocker³ und Hans-Jörg Weiß⁴

1. Anlass für diesen Sammelband

Günter Knieps hat seit seiner Promotion im Jahr 1977 das Forschungsgebiet der Netzökonomie in Deutschland maßgeblich geprägt. Ein in seinen Forschungsarbeiten immer wiederkehrendes Thema ist die Frage nach der „richtigen“ Balance zwischen Wettbewerb und Regulierung in Netzsektoren. Die Entwicklung der Netzindustrien seit ihrer Liberalisierung zeigt, dass dieses „Ausbalancieren“ nie endgültig abgeschlossen ist. Immer wieder tauchen neue Fragestellungen und Herausforderungen auf, die die bis dato erreichte Balance aus dem Gleichgewicht bringen.

Zu Beginn der Liberalisierung stand die Marktmachtproblematik im Zentrum, insbesondere die Frage nach der angemessenen Regulierungsbasis. Bald kamen Probleme aufgrund von Netzexternalitäten dazu, z.B. in Bezug auf die Setzung von Standards. Mit fortschreitender Liberalisierung wurden Allokationsprobleme relevanter, z.B. im Kontext der Bepreisung von Netzinfrastrukturkapazitäten. Heute stellt die Digitalisierung einen zentralen Treiber fundamentaler Herausforderung dar, z.B. im Zusammenhang mit der Weiterentwicklung der Internetarchitektur.

Die oben genannten Entwicklungen in Theorie und Praxis der Netzökonomie spiegeln sich in den Veröffentlichungen von Günter Knieps wider. Unter den vielen wissenschaftlichen Beiträgen, die Günter Knieps bislang vorgelegt hat, genießt ein Beitrag einen besonderen Stellenwert: sein im August 1997 in der Zeitschrift *Kyklos* erschienener Aufsatz „Phasing out Sector-Specific Regulation in Competitive Telecommunications“ (Vol. 50, Issue 3, S. 325–339). Darin hat er seinen disaggregierten Regulierungsansatz erstmals in vollständig ausgearbeiteter Form präsentiert.

1 Pädagogische Hochschule Freiburg; Franziska.Birke@ph-freiburg.de.

2 Jacobs University Bremen; g.brunekreeft@jacobs-university.de.

3 Weizenbaum Institut für die vernetzte Gesellschaft & TU Berlin; vstocker@inet.tu-berlin.de. Volker Stocker bedankt sich für die Unterstützung durch das Bundesministerium für Bildung und Forschung (BMBF) (Förderkennzeichen 16DII124 – „Deutsches Internet-Institut“).

4 Hochschule Offenburg; hans-joerg.weiss@hs-offenburg.de.

Der 25. Jahrestag des Erscheinens dieses Kyklos-Aufsatzes wurde von den Herausgebern zum Anlass genommen, das wissenschaftliche Werk und das Wirken von Günter Knieps als Forscher und Hochschullehrer mit einer Festschrift zu würdigen. Die Herausgeber fühlen sich ihm in besonderer Weise verbunden, da sie alle bei ihm an „seinem“ Institut an der Universität Freiburg i. Br., das er von 1992 bis 2019 geleitet hat, als Assistent(inn)en gearbeitet haben und promoviert wurden bzw. sich habilitiert haben.

Bemerkenswert ist, dass Günter Knieps Forschungsinteressen nie an den Grenzen einzelner Netzindustrien endeten und zudem disziplinäre Grenzen der Netz- und Regulierungsökonomie überschritten. Spektrum und „Offenheit“ seines Schaffens spiegeln sich auch im vorliegenden Sammelband wider. Die Festschrift enthält neben Beiträgen zu Netzindustrien und dort verortbaren Themen auch Beiträge zu Fragestellungen aus diversen Bereichen der Wirtschaftspolitik. Der Autorenkreis umfasst neben Ko-Autoren und Weggefährten aus dem Bereich der netz- und regulierungsökonomischen Forschung auch Kollegen mit anderen wirtschaftspolitischen Forschungsschwerpunkten.

Die nachfolgenden Abschnitte dieses einführenden Beitrags behandeln folgende Themen: Im zweiten Abschnitt werden die Grundzüge des von Günter Knieps entwickelten disaggregierten Regulierungsansatzes skizziert. Im dritten und vierten Abschnitt werden neue Entwicklungen in der Netzökonomik im Energiebereich und Telekommunikation bzw. Internet skizziert. Schließlich werden im letzten Abschnitt die nachfolgenden Beiträge des Sammelbands im Überblick vorgestellt.

2. Der disaggregierte Regulierungsansatz der Netzökonomie⁵

Die Netzindustrien haben in den vergangenen Jahrzehnten in vielen Ländern grundlegende ordnungspolitische Reformen erlebt. So begann z.B. Mitte der 1980er Jahre ein (noch andauernder) Prozess der Liberalisierung der europäischen Verkehrsmärkte. Ein weiteres Beispiel ist die Öffnung der zuvor abgeschotteten und häufig von öffentlichen Monopolunternehmen dominierten Telekommunikationsmärkte. Die bisherigen Erfahrungen mit der Liberalisierung sind überwiegend positiv. Es hat sich jedoch gezeigt, dass auch nach einer Liberalisierung ein signifikanter Regulierungsbedarf in Netzindustrien bestehen bleibt. Für die Ausgestaltung der

5 Wesentliche Teile dieses Abschnitts wurden aus Weiß (2012) übernommen.

Rollenverteilung und Arbeitsteilung zwischen Markt und Staat liefert der disaggregierte Regulierungsansatz eine konzeptionelle Basis. Nachfolgend werden die Grundzüge dieses Ansatzes (in der gebotenen Kürze) skizziert.⁶

Das übergeordnete Ziel staatlicher Regulierung ist die Korrektur von Marktversagenstatbeständen. „Staatliche Regulierungseingriffe sind nur dann allenfalls noch gerechtfertigt, wenn der Wettbewerb und der Markt nicht spontan zu einem gleich guten oder besseren Ergebnis gelangen“ (Knieps, 1997b, S. 256). Die Verfolgung unterschiedlicher sektorspezifischer Regulierungsziele erfordert eine jeweils eigenständige ökonomische Rechtfertigung. Folgende drei grundlegend verschiedene Regulierungsziele stehen in Netzindustrien im Vordergrund: 1.) Marktmachtdisziplinierung und Gewährleistung eines funktionsfähigen Wettbewerbs; 2.) Aufrechterhaltung von politisch erwünschten defizitären Leistungen im Sinne der Daseinsvorsorge; 3.) Sicherstellung technischer Kompatibilität und Ermöglichung Management von gemeinsam genutzten Ressourcen (vgl. Knieps, 1997b, S. 254).

Zentrales Element des disaggregierten Regulierungsansatzes ist das Primat der minimalen Regulierungsbasis, diese „darf nicht über das unbedingt erforderliche Maß ausgedehnt werden (minimaler Regulierungsumfang). Die Gefahr der Überregulierung besteht bei der Verfolgung jedes der drei Regulierungsziele“ (Knieps, 1997b, S. 255). Für die zuverlässige Bestimmung der Regulierungsbasis ist eine Lokalisierung der jeweils regulierungsbedürftigen Netzbereiche per disaggregierter Betrachtung der verschiedenen Netzebenen erforderlich. Hieraus ergibt sich, dass der disaggregierte Regulierungsansatz eine Abkehr von globalen End-zu-End-Regulierungsansätzen bedeutet und solchen beinahe diametral gegenübersteht (Knieps, 2008, S. 96). Zentral für die Bestimmung der Regulierungsbasis ist dabei die Unterscheidung von Netzdiensten und Netzinfrastrukturen (vgl. Knieps, 2007, S. 3).

Dabei ist zu beachten, dass die verschiedenen Netzebenen zueinander vertikal komplementär sind. So erfordert der freie Marktzutritt von Anbietern von Netzdienstleistungen, die keine eigene Netzinfrastruktur besitzen, dass diese einen diskriminierungsfreien Zugang zu den (vorgelagerten) komplementären Netzinfrastrukturkapazitäten erhalten (Knieps,

6 Vgl. hierzu Knieps (1997a, 1997b, 2007, 2008). Um mehrfache Sprachwechsel zu vermeiden, stammen alle direkten Zitate in diesem Abschnitt *nicht* aus dem (englischsprachigen) Kyklos-Aufsatz, sondern aus den genannten drei weiteren (deutschsprachigen) Quellen.

2007, S. 3). Falls nun ein monopolistischer Bottleneck⁷ samt Marktmachtproblem auf der Infrastrukturebene vorliegt, ist eine symmetrische Regulierung erforderlich; so kann die vorliegende Marktmacht diszipliniert werden, und zwar „unabhängig davon, um welches Unternehmen und welche Technologie es sich handelt“ (Knieps, 2008, S. 97). Unabhängig vom netzspezifischen Marktmachtproblem ist eine symmetrische und diskriminierungsfreie Ausübung der technischen Regulierungsfunktionen erforderlich. Diese sind unabdingbar für den Aufbau und Betrieb von Netzen und können auf jeder Netzebene von Bedeutung sein, um die Voraussetzungen für funktionsfähigen Wettbewerb zu schaffen (vgl. Knieps, 2007, S. 135).⁸

Neben dem Erfordernis der disaggregierten Perspektive zur Bestimmung der minimalen Regulierungsbasis, sollten die eingesetzten Regulierungsinstrumente problembezogen und zielspezifisch sein. Dabei ist zu berücksichtigen, „dass es ein perfektes Regulierungsinstrument nicht geben kann und dass Regulierung niemals zu einer perfekten Korrektur des Marktversagens ... führen kann. Es kommen der Einfachheit und der praktischen Implementierbarkeit eine große Bedeutung zu“ (Knieps, 2008, S. 107). Bereits bei Design und Auswahl von Regulierungsinstrumenten ist den Informationsdefiziten und dem begrenzten Wissen der Behörden, die diese Regulierungsinstrumente spezifizieren und später in der Praxis umsetzen sollen, Rechnung zu tragen.

Die eingesetzten Regulierungsinstrumente sollten die unternehmerischen Handlungsspielräume nur so weit einschränken, wie es zur Erreichung des verfolgten Regulierungsziels unbedingt erforderlich ist. Dies kann in manchen Fällen bedeuten, dass die Regulierungsaufgabe von privaten Institutionen wahrgenommen wird. Marktversagen ist nur eine notwendige aber keine hinreichende Bedingung für staatliche Eingriffe. Beispiele für private Verhandlungslösungen finden sich in Netzindustrien z.B. im Bereich der Standardisierung technischer Regulierungsfunktionen durch Komitees (vgl. Knieps, 2007, S. 133f.).

In der Regel erfordert die Umsetzung der Regulierung in Netzindustrien aber die Einrichtung einer Regulierungsbehörde, z.B. in Deutschland der Bundesnetzagentur. Diese sollte möglichst schlank gehalten und unabhängig von der Einflussnahme von Interessengruppen ihre Aufgaben

7 Die Theorie der monopolistischen Bottlenecks, eine Weiter- bzw. Fortentwicklung der Theorie der angreifbaren Märkte, ist ein zentraler Baustein des disaggregierten Regulierungsansatzes (vgl. Knieps, 2006).

8 Zur zielspezifischen Lokalisierung des Regulierungsbedarfs in den Sektoren Eisenbahn und Luftverkehr vgl. Knieps (1996).

wahrnehmen.⁹ Ein zentrales Problem der Regulierungsbehörden in liberalisierten Netzindustrien ist ihre Glaubwürdigkeit im Zeitablauf (Zeitkonsistenzproblem). Um die Selbstbindungsfähigkeit einer Behörde zu stärken, ist ein klar umrissenes Regulierungsmandat seitens des Gesetzgebers notwendig, das den diskretionären Handlungsspielraum der Behörde eindeutig begrenzt (vgl. Knieps, 2007, S. 190–192).

Regulatorisches Mikromanagement widerspricht dem Leitbild einer schlanken Regulierung. Die Regulierungsbehörden sollten Freiräume für Innovationen lassen. So weit wie möglich sollten die Präferenzen der Konsumenten zum Tragen kommen. Insbesondere in dynamischen Kontexten und Netzindustrien kann dies dazu führen, dass die Regulierungsbasis im Zeitablauf schrumpft oder gar vollends wegfällt. Dabei gilt es „[s]taatliche Regulierungseingriffe ... unmittelbar abzuschaffen, wenn der zugrundeliegende Regulierungsbedarf wegfällt (Phasing-out sektorspezifischer Regulierung)“ (Knieps, 1997b, S. 256).

Auch wenn die Netzsektoren auf den ersten Blick sehr unterschiedlich erscheinen, sind die ökonomischen Konzepte im Hintergrund doch sehr ähnlich. Dementsprechend sind die Kernkonzepte des disaggregierten Ansatzes auf verschiedene Netzsektoren anwendbar. Im Grunde geht es immer darum einerseits über eine (vertikale) Umstrukturierung, die Wettbewerbsbereiche zu liberalisieren und andererseits die verbleibenden monopolistischen Bottlenecks sektor-spezifisch zu regulieren. Gleichwohl sind die Details bei den verschiedenen Sektoren sehr unterschiedlich und zum Teil sehr speziell. In seinen Arbeiten ist Günter Knieps dann auch konsequent mehreren Sektoren mit demselben Basiskonzept angegangen und hat damit wesentlich zur Liberalisierung der Verkehrs-, Telekommunikations- und Energiebranchen beigetragen.

Im Folgenden werden ausgehend vom disaggregierten Ansatz neuere Entwicklungen in ausgewählten Netzindustrien – dem Energiesektor sowie dem Telekommunikations- und Internetsektor – skizziert.

3. Neuere Entwicklungen bei den Energienetzen

In den späten 80-er und frühen 90-er Jahren wurden die Energiemärkte liberalisiert und umstrukturiert, Strom, sowie Gas. Man könnte sagen, dass

9 Zur Unterscheidung zwischen einer 'barocken' und einer 'schlanken' Regulierung sowie dem Design einer 'schlanken' Regulierungsbehörde in der Telekommunikation vgl. Knieps (1997b).

die Umstrukturierung weitgehend dem oben dargestellten disaggregierten Ansatz entsprach (vgl. Knieps, 1996). Vertikal wurden die Wettbewerbsbereiche von den monopolistischen Netzen getrennt. Horizontal konnte somit Wettbewerb zugelassen werden bei Produktion, Handel und Vertrieb. Gleichzeitig wurden die monopolistischen Netze aus zwei Gründen reguliert (vgl. Knieps, 2001): Erstens zum Schutz der Netzkunden gegen überhöhte Netztarife und zweitens, um einen diskriminierungsfreien Zugang zu den Netzen für die Wettbewerber zu ermöglichen.

Die Ziele für den Energiebereich bewegen sich immer in einem Zieldreieck: 1. Liberalisierung und Wettbewerb, 2. Versorgungssicherheit und 3. Nachhaltigkeit und Energiewende. Die Entwicklungen und Reformen wurden lange Zeit vom Ziel der Liberalisierung und Wettbewerb geprägt. Das Thema Versorgungssicherheit hat durch den Ukraine Krieg aktuell besondere Bedeutung. In den letzten Jahren allerdings ist das Ziel Nachhaltigkeit und Energiewende dominant geworden. Diese Schwerpunktverschiebung wirkt sich auch auf Marktstruktur und Netzregulierung aus. Unten werden dafür exemplarisch die Entwicklungen der Energienetze in Deutschland dargestellt; jedoch sind diese durchaus vergleichbar mit Erfahrungen in anderen Ländern.

Seit 2009 werden die Erlöse der Energienetze in Deutschland mit Hilfe der sogenannten Anreizregulierungsverordnung (ARegV) reguliert. Die Anreizregulierung ist im Wesentlichen eine Erlösobergrenze (revenue cap) in Anlehnung an der RPI-X-Regulierung wie von Stephen Littlechild (vgl. Littlechild, 1983) entwickelt. Die zentrale Idee ist der Anreiz für kosten-senkende *Effizienzsteigerungen*. Erfahrungen aus dem Ausland und die Erfahrung der ersten Regulierungsperioden in Deutschland lassen vermuten, dass die Anreize effektiv sind. Es wurde aber zwischenzeitlich zunehmend klar, dass diese in einem Konflikt mit kostenerhöhenden *Investitionsanreizen* stehen können (Brunekreeft & Meyer, 2016).

Das Kernproblem liegt im System der Anreizregulierung. Die Anreize für Kosteneffizienz werden gesetzt, indem die regulierten Erlöse erst mit einem fünfjährigen Zeitverzug der Regulierungsperiode an die zu Grunde liegenden Kosten angepasst werden. Bei fallenden Kosten darf der Netzbetreiber den Unterschied während dieser Periode behalten. Umgekehrt gilt aber auch, dass bei steigenden Kosten der Netzbetreiber die zusätzlichen Verluste während der Periode tragen muss. Letzterer Fall tritt bei hohem Investitionsbedarf auf. Die Wirkungsweise des Zeitverzugs in der ARegV wurde in der so genannten „dena-Verteilnetzstudie“ (Dena-VNS, 2012) simuliert: Bei den Netzbetreibern mit hohem Investitionsbedarf bleibt die interne Kapitalverzinsung signifikant hinter der Zielrendite zurück. Nach langer Diskussion wurde mit Start der 3. Regulierungsperiode im Jahr

2019 das System für die Kapitalausgaben geändert: der Zeitverzug wurde aufgehoben und Kapitalausgaben werden jetzt jährlich verrechnet (sog. Kapitalkostenabgleich). Damit machte die preisbasierte Regulierung einen Schritt in Richtung kostenbasierter Regulierung.

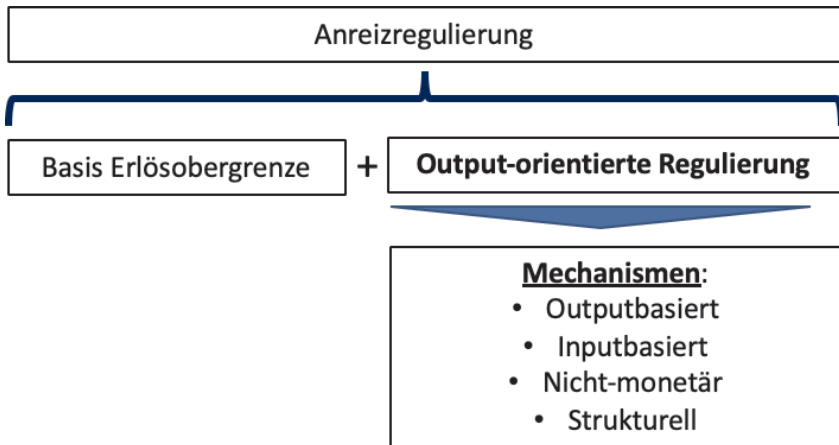
Mittlerweile kündigen sich weitere Entwicklungen an, die sich zwischen Strom- und Gasnetzen stark unterscheiden. Die Entwicklung der Stromnetze ist bedingt von der Energiewende von *Ausbau* geprägt. Die Gasnetze dagegen stehen vor *Rückbau* oder Stilllegung, da die Gasnachfrage abnimmt, oder vor einer Umstellung auf Wasserstofftransport, falls die Wasserstoffentwicklung sich durchsetzt. Diese Entwicklungen erfordern unterschiedliche regulatorische Ausprägungen.

Ein Problem tritt bei der Anreizregulierung dann auf, wenn nicht die Kosten sich ändern, sondern die Nachfrage. Auf der theoretischen Ebene wurde das Thema von Spence (1975, S. 420, Fn. 5) im Kontext der Qualitätsregulierung bei einer kostenbasierten Regulierung behandelt: „*Of somewhat less interest is the case where price is fixed or taken as given. In that case, the firm always sets quality too low.*“ Wenn die Nachfrage sich ändert, entsteht ein *externer* Surplus: *value creation*. Da die Regulierung die Preisentwicklung exogen und ex-ante bestimmt, kann das Unternehmen den zusätzlichen Surplus nicht hinreichend abschöpfen und wird demzufolge ineffizient wenig in Produktverbesserungen investieren.

Insbesondere die Stromnetzbetreiber sehen sich, von der Energiewende angetrieben, zunehmend mit Innovationsbedarf und neuen Aufgaben konfrontiert, die manchmal nur begrenzt mit dem Kerngeschäft verbunden sind.

Hier nun setzt mit output-orientierter Regulierung (OOR) eine neue Entwicklung in der Regulierung an. Output-orientierte Regulierung ergänzt die im Kern fortbestehende Anreizregulierung durch Erlöselemente, die an das Erreichen spezifischer Leistungsziele (Outputs) gekoppelt sind und den gesellschaftlichen Nutzen der Outputs angemessen widerspiegeln sollen. Die bestehende Regulierung wird somit nicht ersetzt, sondern selektiv um einzelne OOR-Elemente ergänzt. Die untere Abbildung 1 fasst die Idee zusammen.

Abbildung 1: Output-orientierte Regulierung



Quelle: Brunekreeft et al. (2020).

Bei den Gasnetzen sehen die Entwicklungen indes anders aus. Die Gasnetze stehen vor der Frage, ob sie stillgelegt, umgebaut oder für eventuelle zukünftige Entwicklungen in Stand gehalten werden. Die Wahl hängt entscheidend von zukünftigen, äußerst unsicheren Entwicklungen ab. Folgende Quellen von Unsicherheit liegen vor:

- Die Entwicklung der Wasserstoffnutzung und -infrastruktur, und damit die Frage, inwiefern die Gasnetze für Wasserstofftransport umgebaut werden sollten,
- die Entwicklung der Gasnachfrage, und
- ein staatlich angeordneter Gasausstieg.

Diese Entwicklung impliziert (zumindest) zwei größere Themen für die zukünftige Regulierung der Gasnetze.

Erstens sollte geprüft werden, inwiefern Realloptionen explizit in der Regulierung berücksichtigt werden sollten. Bei vorliegender Wahl unter Unsicherheit ergibt es viel Sinn, Optionen offen zu halten. Wenn Gasnetze stillgelegt oder gar zurückgebaut werden, verliert der Netzbetreiber die Option später das Netz für Wasserstofftransport umzubauen, sollte dies erwünscht sein. Die Strategie, das Netz in Stand zu halten und zu warten bis klar wird, wie die Zukunft sich entfaltet, ist eine „wait-and-see“ Realloption: Der Netzbetreiber investiert, um die Optionen offen zu halten und zukünftig besser informierte Entscheidungen treffen zu können.

Zweitens entstehen durch den Rückgang der Gasnachfrage u.U. stranded assets im Gasnetz: Es kommt ein Moment, an dem die Netznutzung so gering geworden ist bzw. weiterhin zurückgehen wird, dass Stilllegung besser wäre, auch wenn die Anlagen noch nicht abgeschrieben sind. Demnach stellt sich die Frage, wie solche stranded assets ggf. vermieden werden könnten oder wer die Kosten dieser stranded assets tragen sollte. Hierzu kämen etwa eine Anpassung der Abschreibungsmethodik oder Quersubventionierung z.B. aus dem Stromnetz in Frage.

4. Fundamentale Veränderungen durch Informations- und Kommunikationstechnologien (IKT)

Gegen Ende des letzten Jahrhunderts wurde der Telekommunikationssektor in den meisten Industrienationen umfassend liberalisiert und dereguliert. Die Überführung vormals staatlich geschützter nationaler Monopolunternehmen in ein marktliches Umfeld und die Sicherstellung funktionsfähigen Wettbewerbs erforderte den Einsatz technischer Regulierungen, überdies aber auch und insbesondere die zuverlässige und zielgerichtete Identifikation sowie die wirksame Disziplinierung netzspezifischer Marktmacht (Knieps, 1997a).

In den vergangenen 25 Jahren hat der Telekommunikationssektor eine atemberaubende Entwicklung genommen. Diese setzt sich bis heute fort und hat – insbesondere getrieben durch den digitalen Wandel – die Telekommunikationsindustrie in ein dynamisches Ökosystem verwandelt, das durch eine nie dagewesene Komplexität und Vielfalt an Technologien und Infrastrukturen, Akteuren und Diensten charakterisiert ist (Stocker & Knieps, 2021). Während Entwicklungen im Zusammenhang mit dem Internet der Dinge stellvertretend für eine zunehmende Verschmelzung der realen Welt mit der virtuellen Welt stehen (z.B. Knieps & Bauer, 2016; Knieps, 2017, 2019a, 2019b), sind digitale Infrastrukturen insgesamt und die Kommunikationsinfrastruktur des Internets im Speziellen wesentlich für gesellschaftliche und wirtschaftliche Aktivität, Teilhabe und Resilienz geworden – die Corona-Pandemie hat dies jüngst verdeutlicht (Briglauer & Stocker, 2020; Feldmann et al., 2021; Stocker & Whalley, 2021).

Digitale Transformationsprozesse als Treiber disruptiver Veränderungen

Das Internet-Ökosystem befindet sich in ständigem Wandel und ist durch vielfältige Transformationsprozesse geprägt. So basiert etwa die Bereitstellung von Diensten heutzutage häufig auf einem komplexen und dynamischen Zusammenspiel vernetzter und komplementärer Computing- und Kommunikationsnetzwerkressourcen. Diese Ressourcen sind teilweise im Besitz und unter der Kontrolle verschiedener (oft konkurrierender) Unternehmen; einige der Ressourcen sind offen und werden gemeinsam genutzt, während andere dediziert sind und proprietär betrieben werden (Stocker et al., 2020, 2021).

So hat sich die Art und Weise, wie das Internet zusammengeschaltet ist, wo Daten gespeichert und verarbeitet werden und wie Inhalte und Anwendungen verteilt werden, grundlegend verändert. Althergebrachte Hierarchien und Kräfteverhältnisse der Telekommunikationswelt und des frühen Internets können und dürfen nicht mehr mentalen Modellen zugrunde liegen, die zur Beurteilung der Notwendigkeit sowie dem Design geeigneter Markteingriffen herangezogen werden. Vielmehr müssen Veränderungen vollumfänglich erfasst, mentale Modelle entsprechend angepasst und rückwärtsgerichtete Regulierungsrahmen entsprechend in Frage gestellt werden (Stocker & Knieps, 2021; Stocker et al., 2021).

Eine besondere Rolle in der Evolutorik des Internet-Ökosystems kommt dem Aufschwung sogenannter Big Tech Unternehmen wie Google (Alphabet), Amazon, Netflix oder Facebook (Meta) zu. Während diese Unternehmen in den letzten Jahren in den Auf- und Ausbau privater Netze und Infrastrukturen investiert haben, ist dieser Aufschwung eng verknüpft mit der zunehmenden Bedeutung komplexer, proprietär betriebener Plattformökosysteme. Diese stehen symbolhaft für einen tiefgreifenden Wandel, der sich in fundamental veränderten Wettbewerbs- und Innovationsdynamiken manifestiert und eine teils disruptive Wirkung auf Wertschöpfungsketten, Markt- und Industriestrukturen entfaltet (Lehr et al., 2019; Stocker & Knieps, 2021; Stocker et al., 2021).

Internet und der digitale Wandel haben neue Möglichkeiten zur Automatisierung und Verlagerung gesellschaftlicher und wirtschaftlicher Aktivität in den virtuellen Raum geschaffen bzw. bestehende Möglichkeiten erweitert. Dies zeigt sich auch in Netzindustrien, die zunehmend „smartifiziert“, also intelligenter gemacht, werden (siehe Montero & Finger, 2021). Aus netzökonomischer Perspektive stellen entsprechende Innovationen besonders spannende Phänomene dar. Sie basieren auf einer Verschmelzung der realen/physischen und virtuellen Welt, eröffnen neue Möglichkeiten, bringen jedoch auch neuartige Herausforderungen mit sich – insbesondere

hinsichtlich der ökonomisch effizienten Allokation und Orchestrierung komplementärer Netzressourcen (Knieps & Stocker, 2019; Knieps, 2019a, 2019b, 2021).

Breitband als Schlüsseltechnologie und transformative Kraft

Breitbandnetze spielen eine zentrale Rolle in der Evolutorik des Internet-Ökosystems sowie von Netzindustrien. Während der Wandel von schmalbandigen hin zu breitbandigen Zugangsnetzen sowie deren Weiterentwicklung in puncto Gestalt und Geschwindigkeit maßgeblich von Pfadabhängigkeiten geprägt war und ist (Knieps & Zenhäusern, 2015; Briglauer et al., 2020), zeigt sich ein grundsätzlicher Unterscheid zwischen der traditionellen Telekommunikationswelt und der neuen Welt des breitbandigen Internets.

Herkömmliche Telefonnetze stellten single-purpose Netze dar. Sie waren im Wesentlichen designt, dimensioniert und optimiert für die Bereitstellung von Sprachtelefoniediensten und unterstützten einige wenige weitere Anwendungsdienste. Mit Breitband verhält sich dies grundlegend anders. Zunächst umfasst Breitband eine Vielfalt verschiedener Zugangstechnologien. Diese unterscheiden sich stark hinsichtlich ihrer technologischen Basis (z.B. kabelgebunden per Glasfaser, Kupferdraht oder Hybridtechnologien oder kabellos per Mobilfunk- oder Satellitentechnologie) sowie ihrer Leistungsmerkmale (z.B. bzgl. Datenraten und deren Symmetrie, Latenzzeiten, Performance in Spitzenlastzeiten, Zuverlässigkeit, oder Mobilität). Einzelne Technologien werden weiterentwickelt, neue Technologien werden entwickelt. Moderne, hochbitratige Breitbandnetze (z.B. glasfaserbasierte Zugangstechnologien oder 5G-basierte Mobilfunknetze) unterstützen die (gleichzeitige) Bereitstellung einer sich wandelnden Vielfalt von Anwendungsdiensten (Knieps, 2021; Stocker & Knieps, 2021; Lehr, 2019; Yoo & Lambert, 2019).

Breitbandnetze sind folglich *nicht* für die Bereitstellung eines oder einer kleinen, geschlossenen Menge an Anwendungsdiensten designt, dimensioniert oder optimiert. Vielmehr ermöglichen sie die integrierte Bereitstellung einer sich wandelnden Vielfalt an Anwendungsdiensten. Darüber hinaus unterstützen sie die Entwicklung innovativer Dienste (Knieps & Bauer, 2016; Knieps, 2021; Stocker & Knieps, 2021) und sind Ausgangspunkt einer stark ausgeprägten Innovationsdynamik. Letztere spielt sich sowohl innerhalb als auch zwischen komplementären Netzebenen ab. Innovation auf der nachgelagerten Ebene der Dienste kann Impulse für Innovation (z.B. die Weiterentwicklung) auf vorgelagerten Ebenen geben. Innovation

auf den vorgelagerten Ebenen kann wiederum einen Innovationsimpuls auf der Ebene der Dienste geben, da hierdurch der Möglichkeitenraum für Anwendungsdienste erweitert wird. Aus diesen bidirektionalen Innovationsspillovers kann sich ein dynamisches Zusammenspiel komplementärer Innovationen entwickeln (Bauer & Knieps, 2018).

Entscheidende Voraussetzung für die Entfaltung solcher Dynamiken ist, dass unternehmerische Suchprozesse nach entsprechenden Innovationen nicht ungerechtfertigt (d.h. „künstlich“) eingeschränkt werden. Ungerechtfertigte Regulierungseingriffe können Investitions- und Innovationsanreize verzerren und die Fähigkeit von Betreibern von Breitbandnetzen heterogene Anforderungsprofile eines sich dynamisch ändernden Anwendungsmixes flexibel und ökonomisch effizient zu berücksichtigen einschränken – dies gilt es insbesondere im Kontext von Netzneutralitätsregulierungen zu berücksichtigen (Knieps, 2011b; Bauer & Knieps, 2018; Knieps & Stocker, 2016; Stocker & Knieps, 2019). Ökonomisch effiziente und anreizkompatible Preis- und Qualitätsdifferenzierungsstrategien basieren auf versatilen Verkehrsarchitekturen, die dynamische Priorisierungs- und Kapazitätsreservierungsstrategien ermöglichen und somit die technologische Grundlage für ökonomisch optimale Kapazitätsallokationen bereitstellen. Solche Architekturen werden wesentlich für die effiziente Bereitstellung einer offenen Menge von Anwendungsdiensten, insbesondere wenn sich letztere stark in ihren Anforderungen an den Datentransport (z.B. Kapazitäts- und Qualitätsanforderungen) und die (lokale) Datenverarbeitung (z.B. via Edge Computing/Edge Cloud) sowie hinsichtlich Aspekten wie Mobilität, Energieeffizienz, oder Sicherheit unterscheiden. Schließlich stellen entsprechende Architekturen einen zentralen Baustein für „general-purpose“ Breitbandnetze dar (Knieps, 2015; Knieps & Stocker, 2016; Knieps & Bauer, 2016).

Aktuelle Herausforderungen aus netzökonomischer Perspektive

Aus netzökonomischer Perspektive ergeben sich aus den beschriebenen Entwicklungen eine Reihe von Herausforderungen. In aller Kürze möchten wir im Folgenden zwei dieser Herausforderungen kurz einführend erläutern.

Die Corona-Pandemie hat wie bereits oben erwähnt die Wesentlichkeit von Breitband für gesellschaftliche und wirtschaftliche Resilienz sowie für Teilhabe und Erfolg von Unternehmen und Einzelpersonen verdeutlicht. Um den negativen Folgen digitaler Spaltungen entgegenzuwirken und digitale Inklusion zu fördern, haben insbesondere Fragen der breitbandigen

Mindestversorgung per Universaldienstverpflichtung sowie industriepolitisch motivierte Fragen bezüglich der Förderung hochbitratiger Breitbandnetze eine gesteigerte Relevanz und eine erhöhte Dringlichkeit erlangt. Die Frage „Breitband für alle?“ zu beantworten ist dabei alles andere als trivial. Vielmehr erschwert die oben beschriebene Dynamik und Vielfalt die Bestimmung angemessener Breitbandziele sowie entsprechende Kosten-Nutzen-Abschätzungen.¹⁰ Zunächst ergeben sich eine Reihe von Fragen hinsichtlich der Spezifikation eines angemessenen Versorgungsni-veaus, etwa:

- Welche (Anwendungs-)Dienste sollen ermöglicht werden (Frage nach dem „Dienstekorb“)?
- Welche Leistungsmerkmale lassen sich aus dem „Dienstekorb“ für die Spezifikation eines angemessenen Zugangsdienstes ableiten?
- Wie kann der Dynamik ständig steigender Anforderungen angemessen Rechnung getragen werden (z.B. [wie] können oder sollten Versorgungsziele dynamisiert werden?)

In der Folge ergeben sich zudem Fragen bezüglich der Durchsetzung und Zielerreichung, etwa:

- Wie kann im Rahmen entsprechender Förderstrategien Technologie-neutralität gewahrt werden?
- Welche Formen des Netzmanagements sind notwendig, um die Versorgungsziele zu erreichen?
- Wie kann sichergestellt werden, dass Unternehmen und Einzelpersonen entsprechende Angebote wahrnehmen und zu ihrem Vorteil nutzen können?

Diese und weitere Fragestellungen prägen seit geraumer Zeit – und seit Ausbruch der Corona-Pandemie vermehrt – Diskussionen in der EU sowie in den USA.¹¹

10 Siehe hierzu frühe Diskussionen zum breitbandigen Universaldienst in Birke (2009a, 2009b) sowie die Diskussion des breitbandigen Universaldienstes im Vereinigten Königreich in Stocker und Whalley (2019). Eine Diskussion der Verankerung eines breitbandigen Universaldienstes in Richtlinie (EU) 2018/1972 des Europäischen Parlaments und des Rates findet sich in Stocker (2020). Briglauer und Stocker (2020) diskutieren die Wesentlichkeit von Breitband für gesellschaftliche und wirtschaftliche Resilienz im Kontext der Corona-Pandemie.

11 In der EU sind gegenwärtig Mitgliedsstaaten mit der Umsetzung der in Richtlinie (EU) 2018/1972 formulierten Universaldienstziele befasst. Zudem wurde im Rahmen der Initiative *Europas Digitale Dekade* jüngst das Ziel „Gigabit-Netzanbindung für alle europäischen Haushalte und 5G-Versorgung aller besiedelten Gebie-

Eng verknüpft mit Fragen des Ausbaus breitbandiger Zugangsnetze sind Fragen nach der optimalen Nutzung bestehender Kapazitäten. Wie kann eine sich wandelnde Vielfalt an Anwendungsdiensten und Use Cases, die sich stark hinsichtlich ihrer Anforderungsprofile an den Datentransport sowie in Bezug auf cloud-basierte (lokale) Datenverarbeitung unterscheiden, integriert und in einer ökonomisch effizienten Weise bereitgestellt werden? Günter Knieps hat sich in seiner Forschung der vergangenen Jahre intensiv mit dieser und angrenzenden Fragestellungen befasst. So hat er früh auf die Notwendigkeit innovativer und flexibler Netzarchitekturen zur Implementierung anreizkompatibler Preis- und Qualitätsdifferenzierungsstrategien beim Datentransport im Internet zur Erzielung ökonomisch effizienter Kapazitätsallokationen hingewiesen und durch entsprechende modelltheoretische Ansätze fundiert (Knieps, 2011a, 2011b, 2015). Aufbauend auf den gewonnenen Erkenntnissen hat er in diesem Zusammenhang Spannungsverhältnisse mit starren Netzneutralitätsregulierungen aufgezeigt und für ein marktgetriebenes und adaptives Verständnis von Netzneutralität plädiert (Knieps, 2011a; Knieps & Stocker, 2015, 2016).

Mit der Weiterentwicklung hin zu modernen glasfaser- oder 5G-basierten Breitbandnetzen und der damit einhergehenden rasant wachsenden Vielfalt an Anwendungsdiensten und Use Cases wird dieses Spannungsverhältnis weiter verschärft – Allokationsprobleme gewinnen nicht nur an Komplexität, sondern auch an Bedeutung (Stocker & Knieps, 2019, 2021; Yoo & Lambert, 2019; Fetzer, 2019). Zur Bewältigung und Auflösung solcher Probleme, werden virtuelle Netze wichtiger. Basierend auf der Kombination komplementärer (gemeinsam genutzter) physischer Ressourcen (z.B. Bandweitenkapazitäten, Sensornetze, oder (lokale) Cloudkapazitäten) können solche Netze flexibel erstellt und dynamisch angepasst werden. Virtuelle Netze unterstützen die effiziente Bereitstellung verschiedenster Klassen von Anwendungsdiensten und Use Cases und sind in der Lage deren heterogene Anforderungsprofile maßgeschneidert zu bedienen (Knieps, 2019, 2021; Knieps & Stocker, 2019). In Form von sogenannten ‚network slicing‘ Ansätzen stellen solche Netze bereits Kernelemente von 5G und 6G Mobilfunknetzen dar – die immensen Potenziale dieser Tech-

te“ verankert – dieses soll bis 2030 erreicht werden (Europäische Kommission, 2021, S. 7). In den USA wurde 2021 der Infrastructure Investment and Jobs Act verabschiedet; dieser sieht 65 Mrd. \$ für Breitbandförderung zur Schließung digitaler Spaltungen vor (The White House, 2021).

nologien sollten nicht ungerechtfertigt durch rückwärtsgerichtete Regulierungsansätze (z.B. starre Netzneutralitätsregeln) unausgeschöpft bleiben.

5. Vorschau auf die nachfolgenden Beiträge

Die vorliegende Festschrift beinhaltet zehn Beiträge, die sich mit diversen netzökonomischen oder wettbewerbs- oder wirtschaftspolitischen Fragen beschäftigen. Im Folgenden geben wir einen kleinen Ausblick über die jeweiligen Beiträge. Kriterium für die Bestimmung der Reihenfolge der Beiträge war dabei nicht etwa ein thematisches Clustering; vielmehr wurden die Beiträge entlang der Nachnamen der Autoren alphabetisch sortiert.

(1) Johannes Bauer: „Agile Regulation: Experiments in Flexible Governance“

Mit einem Fokus auf Aspekten der Netzinfrastrukturpolitik untersucht der Beitrag die Rolle agiler und anderer Formen flexibler Regulierung im Informations- und Kommunikationssektor. Es wird aufgezeigt, dass alle Formen von Governance als dynamische, lernende Systeme betrachtet werden müssen. Grundvoraussetzung für agile Regulierungsformen und flexible Governancesysteme sind (i) die Verknüpfung eines wirksamen Monitoringprozesses mit systematischem Lernen sowie (ii) die Generierung von Feedback, das geeignete Anpassungen der Governance vorschlägt.

(2) Falk von Bornstaedt: „Preisbildung bei der Zusammenschaltung im Internet“

Falk von Bornstaedt setzt sich in seinem Beitrag kritisch mit dem fundamentalen Wandel des Internets hin zu einer videozentrierten Distributionsplattform sowie dessen Konsequenzen und Implikation bezüglich Verhandlungsmacht und Kräftegleichgewicht im Internet-Ökosystem auseinander. Ein zentrales Augenmerk gilt der Analyse vorherrschender Bepreisungsmechanismen im Kontext von Zusammenschaltungen im Internet, insbesondere der sogenannten „95th percentile rule“. Während Unzulänglichkeiten und Fehlanreize identifiziert werden, endet der Beitrag mit einem Plädoyer für eine Neubewertung etablierter Bepreisungsmechanismen und einen Wechsel hin zu innovativen, anreizkompatiblen Bepreisungsstrategien.

(3) Manfred J. Holler und Florian Rupp: “Networks, Centrality, Public Good Index and the Medici”

Dieser Beitrag diskutiert einen innovativen Ansatz zur Analyse der Rolle und Bedeutung von Macht in Netzwerken. Der Ansatz vermeidet den

Umweg über die alternative Verwendung graphentheoretischer Maße zur Erfassung der Zentralität in Netzwerken und wendet stattdessen den Public Good Index an. Auf dieser konzeptionellen Grundlage wird das historische Beispiel der Familie Medici im Florenz des 15. Jahrhunderts untersucht. Das von den Medici durch verschiedene Heiraten geschaffene Familien-Netzwerk wird im Hinblick auf die sich daraus ergebende Rolle und Bedeutung von Macht analysiert. Es wird u.a. aufgezeigt, dass dies ein wichtiger Erfolgsfaktor bzw. „Treiber“ des politischen Erfolgs von Cosimo de' Medici war. Der Beitrag schließt mit einigen Bemerkungen zur Heterogenität aufgrund von Netzwerkbeziehungen und diskutiert einige Gemeinsamkeiten zwischen räumlichen Modellen zur Analyse wirtschaftlicher Aktivitäten und räumlichen Modellen zur Analyse politischer Aktivitäten.

(4) *Hans-Ulrich Küpper: „Lessons learned – Anstöße der Regulierung zum Abschreibungsproblem“*

In seinem Beitrag erörtert Hans-Ulrich Küpper die Entwicklungen der Abschreibungsproblematik in der Regulierung. Gemeinsam mit Günter Knieps wurde Ende der 90er Jahre im Hinblick auf die Ermittlung kostenorientierter Netzentgelte das Abschreibungsproblem neu analysiert. Der Prozess der Regulierung von Netzindustrien hat weltweit zu neuen Einsichten geführt. Die erneute Analyse des Abschreibungsproblems deutet auf einen grundlegenden Aspekt hin, der sich bei ihrer Bestimmung für unternehmerische Entscheidungsprobleme noch klarer gezeigt hat. Maßgeblich für ihre Bestimmung sind weniger vergangene als die gegenwärtigen und künftigen Anlagenwerte. Soweit Abschreibungen als Informationen für (externe) Entscheidungen in der Regulierung über künftige Preise ihrer Nutzung oder für innerbetriebliche Entscheidungen über ihre Nutzung bereitzustellen sind, hat sich der Blick grundsätzlich zu ändern. Ihre Bestimmung sollte weniger auf vergangene Anschaffungs- als auf zukünftige Gebraucht- und Anschaffungswerte blicken.

(5) *Kay Mitusch: „Der Deutschland-Takt: Regulierung, Wettbewerb und Organisation auf der Schiene“*

In diesem Beitrag setzt sich Kay Mitusch kritisch mit zu Grunde liegenden organisatorischen und wettbewerbspolitischen Fragen auseinander. Ein wesentlicher Teil des Beitrags exploriert und erörtert aus netzökonomischer Perspektive Möglichkeiten, die sich im Zusammenhang mit dem „Wettbewerb im und zum Deutschland-Takt“ ergeben. Darüber hinaus werden Fragen der Finanzierung und damit einhergehende anreiz- und

politökonomische Aspekte kritisch diskutiert. Schließlich werden fünf Vorschläge für die Zukunft des Deutschland-Takts unterbreitet.

- (6) *Friedrich Schneider: „Der Beitrag von (nicht-)ökonomischen Faktoren zur Analyse von Korruption, Steuerhinterziehung, Steuermoral und Schwarzarbeit“*

In dieser Arbeit werden zunächst einige Ausführungen über die theoretische Modellierung und empirische Evidenz zu wesentlichen Erklärungsfaktoren über Schwarzarbeit und Korruption gegeben. Als klassische Ursachen von Schwarzarbeit gelten die Belastung mit Steuern, Sozialversicherungsbeiträge auf den Produktionsfaktor Arbeit, eine intensive Regulierung, ein unzureichendes und/oder qualitativ schlechtes Angebot an staatlichen Gütern und Dienstleistungen, eine schwache Steuermoral, geringe Rechtssicherheit und nicht funktionierende öffentliche Institutionen, wobei Korruption auch wesentlich von den letzten drei Faktoren verursacht wird. Außerdem werden einige Schätzverfahren und dann empirische Ergebnisse über die Größenordnung von Schwarzarbeit vorgestellt. Anschließend werden aus der Mikroperspektive die (nicht-)ökonomischen Bestimmungsgründe der Steuermoral erörtert, z.B. das Ausmaß der Strafen, die Bildung und das sozialpolitische Engagement. Um Schwarzarbeit und Korruption einzudämmen, sollte der Staat neben Strafen ein ausreichendes und qualitativ hochstehendes Angebot an öffentlichen Gütern und Dienstleistungen bereitstellen.

- (7) *Viktor J. Vanberg: „Wettbewerbspolitik und Demokratie: Das Spannungsfeld von Konsumenten-, Produzenten- und Bürgerinteressen“*

Viktor J. Vanberg erweitert in seinem Beitrag das Spannungsfeld zwischen Produzenten- und Konsumenteninteressen, das bei der Legitimation eines Leitbildes der Wettbewerbspolitik in den Blick genommen werden muss, um die Bürgerinteressen. Während die Konsumentensouveränität als Ordnungsideal der Wettbewerbspolitik angesehen werden kann, weil es im gemeinsamen Interesse der Bürger liegt, ihren Konsumenteninteressen gegenüber (ihren) protektionistischen Produzenteninteressen Vorrang einzuräumen, können globalisierungskritischen Strömungen in dieser Gegenüberstellung nur unzureichend erklärt werden. Vielmehr müssen Bürger bei der Gestaltung der wirtschaftlichen Ordnung ihres Gemeinwesens zudem auch zwischen ihrem Interesse, als Konsumenten die Vorzüge des Zugangs zu einem großen Markt nutzen zu können, und ihrem Interesse daran, gemeinsam mit ihren Mitbürgern über Dinge entscheiden zu können, die für die Lebensqualität, die ihnen ihr Gemeinwesen bietet, von Bedeutung sind, abwägen.

(8) *Achim Wambach: „Die Anti-Stau-Maut“*

Achim Wambach nimmt das 2014 vom Wissenschaftlichen Beirat beim Bundesministerium für Wirtschaft und Energie unter Federführung von Günter Knieps erstellte Gutachten zu „Engpassbasierte[r] Nutzerfinanzierung und Infrastrukturinvestitionen in Netzsektoren“ in den Fokus. In diesem Gutachten empfahl der Beirat verkehrsabhängige Straßennutzungsgebühren, regionale Preise im Stromnetz, und qualitäts-differenzierende Preise der Internetnutzung. Achim Wambach geht in seinem Beitrag der Frage nach, was aus diesen Empfehlungen in den Sektoren Verkehr, Elektrizität und Telekommunikation & Internet geworden ist.

(9) *Bernhard Wieland: „Der disaggregierte Regulierungsansatz und die preußischen Eisenbahnen im 19. Jahrhundert“*

In seinem Beitrag macht Bernhard Wieland einen wirtschaftsgeschichtlichen Ausflug in den disaggregierte Regulierungsansatz und die preußischen Eisenbahnen im 19. Jahrhundert. Die preußische Eisenbahngeschichte wird bisweilen als Beleg dafür angesehen, dass Wettbewerb im Eisenbahnwesen nicht funktionieren könne. Fremdling und Knieps (1993) haben diese These einer eingehenden Analyse unterzogen. Wielands Beitrag ergänzt diese Analyse in zweierlei Hinsicht. Zum einen wird eine Schrift von David Hansemann aus dem Jahr 1837 erörtert, die überraschende Ähnlichkeiten zu Knieps' Ansatz der Disaggregierten Regulierung aufweist. Zum anderen wird auf die von Fremdling und Knieps vernachlässigte Rolle Bismarcks in der preußischen Eisenbahnpolitik eingegangen.

(10) *Patrick Zenhäusern: „Access Regulation of Non-personal Data Markets“*

In seinem Beitrag untersucht Patrick Zenhäusern Fragen bezüglich des Zugangs zu nicht-personenbezogenen Daten aus Wettbewerbs- und regulierungsökonomischer Perspektive. Marktversagenstatbestände und die Dynamik entsprechender Datenmärkte werden erörtert und die Relevanz von Investitionsanreizen erläutert. Bei der Analyse des Interventionsbedarfs zur Sicherstellung des diskriminierungsfreien Zugangs zu nicht-personenbezogene Daten wird die Arbeitsteilung zwischen der Anwendung des allgemeinen Wettbewerbsrechts (ex post) und sektorspezifischer Regulierung (ex ante) diskutiert. Es wird vorgeschlagen, dass Marktuntersuchungen („market surveys“/„market investigations“) als zusätzliches Instrument des Wettbewerbsrechts nach dem Vorbild der britischen und niederländischen Wettbewerbsbehörden geprüft werden und betont, dass nur bei einem systematischen Vorliegen und erwartbarem Missbrauch von

Marktmacht im Rahmen des Datenzugangs Bedarf an sektorspezifischen ex ante Regulierungen (z.B. per verpflichtender Lizenzvergabe) besteht.

Referenzen

- Bauer, J. M., & Knieps, G. (2018). Complementary innovation and network neutrality. *Telecommunications Policy*, 42(2), 172–183. <https://doi.org/10.1016/j.telpol.2017.11.006>.
- Birke, F. (2009a). Universaldienst in Breitbandnetzen. In G. Knieps (Hrsg.). *Fallstudien zur Netzökonomie* (pp. 53–86). Wiesbaden: Gabler.
- Birke, F. (2009b). *Zum Wandel des Universaldienstes in der Telekommunikation: Eine netzökonomische Analyse*. Baden-Baden: Nomos.
- Briglauer, W., & Stocker, V. (2020). *Bedeutung digitaler Infrastrukturen und Dienste und Maßnahmen zur Förderung der Resilienz in Krisenzeiten* (Policy Note No. 42). EcoAustria. <https://www.econstor.eu/handle/10419/227422>.
- Briglauer, W., Stocker, V., & Whalley, J. (2020). Public policy targets in EU broadband markets: The role of technological neutrality. *Telecommunications Policy*, 44(5), 101908. <https://doi.org/10.1016/j.telpol.2019.101908>.
- Brunekreeft, G., Kuszniir, J., & Meyer, R. (2020). *Output-orientierte Regulierung—Ein Überblick* (Working Paper No. 35). Bremen Energy Working Papers. <https://www.econstor.eu/handle/10419/247215>.
- Brunekreeft, G., & Meyer, R. (2016). Anreizregulierung bei Stromverteilnetzen: Effizienz versus Investitionen. *Perspektiven Der Wirtschaftspolitik*, 17(2), 172–187. <https://doi.org/10.1515/pwp-2016-0010>.
- Dena-VNS (2012). *Ausbau- und Innovationsbedarf der Stromverteilnetze in Deutschland bis 2030* (Regulatorisches Gutachten; Dena-Verteilnetzstudie). Deutsche Energie-Agentur GmbH (dena).
- Europäische Kommission (2021). *Digitaler Kompass 2030: der europäische Weg in die digitale Dekade* COM(2021) 118 final, 9. März, Brüssel.
- Feldmann, A., Gasser, O., Lichtblau, F., Pujol, E., Poese, I., Dietzel, C., Wagner, D., Wichtlhuber, M., Tapiador, J., Vallina-Rodriguez, N., Hohlfeld, O., & Smaragdakis, G. (2021). A year in lockdown: How the waves of COVID-19 impact internet traffic. *Communications of the ACM*, 64(7), 101–108. <https://doi.org/10.1145/3465212>.
- Fetzer, T. (2019). Net Neutrality in Europe after the Net Neutrality Regulation 2015/2120. In G. Knieps & V. Stocker (Hrsg.). *The Future of the Internet: Innovation, Integration and Sustainability*. Baden-Baden: Nomos. <https://doi.org/10.5771/9783748902096>.
- Fremdling, R., & Knieps, G. (1993). Competition, regulation and nationalization: The Prussian railway system in the nineteenth century. *Scandinavian Economic History Review*, 41(2), 129–154. <https://doi.org/10.1080/03585522.1993.10415864>.
- Knieps, G. (1996). *Wettbewerb in Netzen: Reformpotentiale in den Sektoren Eisenbahn und Luftverkehr*. Tübingen: Mohr.

- Knieps, G. (1997a). Phasing out Sector-Specific Regulation in Competitive Telecommunications. *Kyklos*, 50(3), 325–339. <https://doi.org/10.1111/1467-6435.00019>.
- Knieps, G. (1997b). Ansätze für eine “schlanke” Regulierungsbehörde für Post und Telekommunikation in Deutschland. *ORDO*, 48, 253–268. <https://doi.org/10.1515/9783110505344-016>
- Knieps, G. (2001). *Wettbewerbsökonomie: Regulierungstheorie, Industrieökonomie, Wettbewerbspolitik*. Berlin & Heidelberg: Springer.
- Knieps, G. (2006). Von der Theorie angreifbarer Märkte zur Theorie monopolistischer Bottlenecks. In F. Daumann, S. Okruch, & C. Mantzavinos (Hrsg.). *Wettbewerb und Gesundheitswesen: Konzeptionen und Felder ordnungsökonomischen Wirkens, Festschrift für Peter Oberender zu seinem 65. Geburtstag* (pp. 141–159). Budapest: Andrásy Gyula Deutschsprachige Universität.
- Knieps, G. (2007). *Netzökonomie: Grundlagen, Strategien, Wettbewerbspolitik*. Wiesbaden: Gabler.
- Knieps, G. (2008). *Wettbewerbsökonomie: Regulierungstheorie, Industrieökonomie, Wettbewerbspolitik* (3., durchges. und aktualisierte Aufl.). Berlin & Heidelberg: Springer.
- Knieps, G. (2011a). Market Driven Network Neutrality and the Fallacies of Internet Traffic Quality Regulation. *International Telecommunications Policy Review*, 18(3), 1–22.
- Knieps, G. (2011b). Network neutrality and the evolution of the internet. *International Journal of Management and Network Economics*, 2(1), 24–38. <https://doi.org/10.1504/IJMNE.2011.042578>.
- Knieps, G. (2015). Entrepreneurial traffic management and the Internet Engineering Task Force. *Journal of Competition Law and Economics*, 11(3), 727–745. <https://doi.org/10.1093/joclec/nhv018>.
- Knieps, G. (2017). Internet of Things, future networks, and the economics of virtual networks. *Competition and Regulation in Network Industries*, 18(3–4), 240–255. <https://doi.org/10.1177/1783591718784398>.
- Knieps, G. (2019a). Internet of Things (IoT), heterogeneous virtual networks and the future of the Internet. In G. Knieps & V. Stocker (Hrsg.). *The Future of the Internet: Innovation, Integration and Sustainability* (pp. 19–32). Baden-Baden: Nomos. <https://doi.org/10.5771/9783748902096>.
- Knieps, G. (2019b). Internet of Things, big data and the economics of networked vehicles. *Telecommunications Policy*, 43(2), 171–181. <https://doi.org/10.1016/j.telpol.2018.09.002>.
- Knieps, G. (2021). Digitalization technologies: The evolution of smart networks. In J. Montero & M. Finger (Hrsg.). *A Modern Guide to the Digitalization of Infrastructure* (pp. 43–58). Cheltenham: Edward Elgar Publishing. <https://doi.org/10.4337/9781839106057.00008>.

- Knieps, G., & Bauer, J. M. (2016). The industrial organization of the Internet. In J. M. Bauer & M. Latzer (Hrsg.). *Handbook on the Economics of the Internet* (pp. 23–54). Cheltenham: Edward Elgar Publishing. https://ideas.repec.org/h/elg/eech ap/14700_2.html
- Knieps, G., & Stocker, V. (2015). Network Neutrality Regulation: The Fallacies of Regulatory Market Splits. *Intereconomics*, 50(1), 46–51. <https://doi.org/10.1007/s10272-015-0524-4>.
- Knieps, G., & Stocker, V. (2016). Price and QoS differentiation in all-IP networks. *International Journal of Management and Network Economics*, 3(4), 317–335. <https://doi.org/10.1504/IJMNE.2016.079877>.
- Knieps, G., & Stocker, V. (2019). Introduction. In G. Knieps & V. Stocker (Hrsg.). *The Future of the Internet: Innovation, Integration and Sustainability* (pp. 7–10). Baden-Baden: Nomos. <https://doi.org/10.5771/9783748902096-7>.
- Knieps, G., & Zenhäusern, P. (2015). Broadband Network Evolution and Path Dependency. *Competition and Regulation in Network Industries*, 16(4), 335–353. <https://doi.org/10.1177/178359171501600402>.
- Lehr, W. (2019). 5G and the Future of Broadband. In G. Knieps & V. Stocker (Hrsg.). *The Future of the Internet: Innovation, Integration and Sustainability* (pp. 109–150). Baden-Baden: Nomos. <https://doi.org/10.5771/9783748902096-109>.
- Lehr, W., Clark, D., Bauer, S., Berger, A., & Richter, P. (2019). Whither the Public Internet? *Journal of Information Policy*, 9, 1–42. <https://doi.org/10.5325/jinfopoli.9.2019.0001>.
- Littlechild, S. C. (1983). *Regulation of British Telecommunications' Profitability*. London: Great Britain Department of Industry.
- Montero, J., & Finger, M. (2021). *The Rise of the New Network Industries: Regulating Digital Platforms*. New York, NY: Routledge. <https://doi.org/10.4324/9781003141327>.
- Spence, A. M. (1975). Monopoly, Quality, and Regulation. *The Bell Journal of Economics*, 6(2), 417–429. <https://doi.org/10.2307/3003237>.
- Stocker, V. (2020). *Innovative Capacity Allocations for All-IP Networks: A Network Economic Analysis of Evolution and Competition in the Internet Ecosystem*. Baden-Baden: Nomos.
- Stocker, V., & Knieps, G. (2019). Network Neutrality Through the Lens of Network Economics. *Review of Network Economics*, 17(3), 115–150. <https://doi.org/10.1515/rne-2019-0025>.
- Stocker, V., & Knieps, G. (2021). Digitalizing telecommunications: Innovation, complexity and diversity in the internet ecosystem. In J. Montero & M. Finger (Hrsg.). *A Modern Guide to the Digitalization of Infrastructure* (pp. 59–91). Cheltenham: Edward Elgar Publishing. <https://doi.org/10.4337/9781839106057.00009>.
- Stocker, V., Knieps, G., & Dietzel, C. (2021). *The Rise and Evolution of Clouds and Private Networks – Internet Interconnection, Ecosystem Fragmentation*. <https://doi.org/10.2139/ssrn.3910108>.

- Stocker, V., Smaragdakis, G., & Lehr, W. (2020). The state of network neutrality regulation. *ACM SIGCOMM Computer Communication Review*, 50(1), 45–59. <https://doi.org/10.1145/3390251.3390258>.
- Stocker, V., & Whalley, J. (2019). Who replies to consultations, and what do they say? The case of broadband universal service in the UK. *Telecommunications Policy*, 43(8), 101823. <https://doi.org/10.1016/j.telpol.2019.04.007>.
- Stocker, V., & Whalley, J. (2021). *The Internet has coped well with Covid-19, but problems remain: Evidence to House of Lords Committee exploring the impact of Covid-19* (Position Paper; Weizenbaum Series 15). Weizenbaum Institute for the Networked Society. <https://www.ssoar.info/ssoar/handle/document/72996>.
- The White House (2021). *UPDATED FACT SHEET: Bipartisan Infrastructure Investment and Jobs Act*. <https://www.whitehouse.gov/briefing-room/statements-releases/2021/08/02/updated-fact-sheet-bipartisan-infrastructure-investment-and-jobs-act/>.
- Weiß, H.-J. (2012). Die traditionelle Bankenregulierung aus Sicht des disaggregierten Regulierungsansatzes der Netzökonomie. In M. Otte & T. Seyfriedt (Hrsg.), *Management 2012: Kurzbeiträge aus Wissenschaft und Praxis* (pp. 107–127). Charleston, SC: CreateSpace.
- Yoo, C. S., & Lambert, J. (2019). 5G and Net Neutrality. In G. Knieps & V. Stocker (Hrsg.), *The Future of the Internet: Innovation, Integration and Sustainability* (pp. 221–245). Baden-Baden: Nomos. <https://doi.org/10.5771/9783748902096-221>.

Agile Regulation: Experiments in Flexible Governance

Johannes M. Bauer¹

Abstract

The Organisation for Economic Co-operation and Development (OECD) and national policymakers are increasingly embracing agile regulation. They see it as a response to the new patterns of scientific and technological innovation often referred to as the Fourth Industrial Revolution. Agile regulation uses instruments, such as anticipatory regulation, outcome-focused regulation, data-driven regulation, and experimental regulation to support fast-paced, cross-sectoral innovation while mitigating the potential harms of such rapid change. These methods are seen as complements and substitutes to traditional regulation, which is perceived as comparatively rigid and slow. Despite the heightened interest in agility, the notion that regulation sometimes needs to adapt quickly is not entirely new. With a focus on network infrastructure policy, this chapter examines the past and future roles of agile forms of regulation in the digital economy. In a broader historical arc, agile regulation is but the latest adaptation of governance in the co-evolution of sector conditions and policy approaches. It is primarily a practical response to improve the procedural rationality of governance, but is built on weaker theoretical foundations than traditional regulation. It will therefore be important to establish appropriate learning and knowledge-sharing frameworks that will contribute to a gradual refinement of the approach.

Keywords:

agile regulation, adaptive regulation, digital economy, dynamic systems, uncertainty

1 Professor, Department of Media and Information; Director, Quello Center, Michigan State University; bauerj@msu.edu. The author thanks Volker Stocker for valuable comments. The chapter is dedicated to Günter Knieps, with gratitude for years of invigorating discussions and the hope for continued collaboration.

1. Introduction

This chapter examines the contributions of agile forms of regulation to surmount pressing governance challenges of the digital economy. A consensus is growing among stakeholders that harnessing the benefits of new digital technologies requires different, innovative public policies. Agile regulation promises novel approaches that align well with the conditions of rapidly advancing technology. However, the discussion is only at the beginning, and little is known about the strengths and limitations of agile regulation. Could it offer effective means to neutralize the market power of dominant players, such as Google, Facebook, and Amazon? Could it offer governance approaches to maintain secure and reliable operations of cyber-physical systems, the Internet of Things (IoT), and Artificial Intelligence (AI)? Might agile regulation offer new tools to mitigate the growing concerns about harms from misinformation and disinformation?

The Fourth Industrial Revolution (Schwab, 2016) coincides with an acceleration of scientific and technological change. At the same time, change increasingly transcends industrial sectors and national borders. Traditional forms of regulation are often too slow and rigid to respond effectively to rapid processes of transformation. Several recent reports by the World Economic Forum (WEF) and the OECD posit that flexible and agile governance is imperative to unleash this scientific and innovation potential (OECD, 2021a, 2021b; World Economic Forum, 2020). Agility is considered vital to address four challenges of the regulation of cross-sectoral, fast-paced change: the “pacing problem” of adapting the speed of regulation to the speed of technological change, the alignment of regulation with the techno-economic conditions, regulatory enforcement, and the institutional and transboundary challenges of change (OECD, 2021b).

It is not necessarily new that public policy and regulation need to be nimble and adaptive. Because problems change over time, policies typically adapt, even though the rate of change may be glacial. The notion of adaptive governance emerged in the 1990s in the context of social-ecological systems (Folke, Hahn, Olsson, & Norberg, 2005; Steelman, 2016). Subsequently, the concept was applied to other areas, including governance of the Internet (Brass & Sowell, 2020; Brunner, 2010; Cherry & Bauer, 2004). Agile regulation is a newer and less developed concept that emerged in response to more recent changes in the pattern of scientific progress in nanotechnology, biotechnology, and the digital economy (OECD, 2021b; World Economic Forum, 2020, p. 6). These new forms of governance are envisioned to support ubiquitous technological and economic change

better than traditional methods of regulation. They are also considered superior to mitigate some of the potential harms of recent technologies.

At a first glance, the ideas and metaphors behind agile regulation are appealing and timely: fast-paced, global technological innovation is met with similarly flexible and unbureaucratic public policy. Many of the practical details of applying these concepts in the current environment remain to be worked out. Only sustained practice will show whether the expectations can be realized. The goal of this chapter is to examine the promises and limitations of agile regulation with a focus on information and communication technologies (ICTs). Among other issues, it will explore whether agility and the desire for regulatory stability can be reconciled, the risk of creating new forms of efficiency-reducing transaction costs, challenges related to the political process of regulation, and the theoretical and normative foundations of agile regulation.

The next section positions the current discussion in the broader historical context of efforts to adapt regulation to the material conditions of the ICT industries. Section three discusses the role of flexible regulation in aligning governance goals with the changing technical and economic conditions of advanced industrial sectors. The potential role of agile forms of regulation in the information and communications sector is reviewed in section four. Section five critically examines the pros and cons of agile regulation and sketches viable alternative paths forward. The main arguments of the chapter are summarized in the conclusion.

2. From traditional to flexible governance

Adaptive and agile regulation are pragmatic responses to the perceived inadequacies and challenges of prevailing forms of regulation. First, they seek to adapt governance to the faster rhythms of the generation of scientific knowledge, technological change, and innovation. Second, they seek to reduce known shortcomings of regulation, such as problems of asymmetric information. Third, they are a reaction to broader transformations in the digital economy, such as the intensifying interdependencies among firms and nations, the global nature of many innovations, and the increased uncertainty they entail for national policy.

At a fundamental level, all regulation contains elements of adaptivity. In North America, regulatory agencies were established in the late nineteenth century, because the legislative process was perceived to be too slow and ineffective to address the highly specialized issues raised by transportation, energy, and telecommunications. Later, environmental regulations, health

and safety regulations, and a growing range of interventions to mitigate the risk of modern technologies were added. Outside of North America, infrastructure industries were typically organized as state-owned enterprises that combined regulatory and operational roles in one entity. With (partial) privatization and market reforms starting in the 1970s, an increasing number of countries established independent regulatory agencies as a more flexible and efficient institutional model. Environmental and other forms of regulation likewise proliferated (Ayres & Braithwaite, 1992; Majone, 1996; Viscusi, Harrington, & Sappington, 2018).

Regulation by specialized government agencies was overhauled starting in the 1960s. While the United States embarked on a process of regulatory reform and deregulation, the model of independent regulation was emulated and adapted in an increasing number of countries looking for alternatives to the historical governmental monopoly in the sector. Despite this streamlining and rejuvenation, a keen sense prevailed that the traditional approach to regulation impeded technological advancements in networking and communications. In response, the Internet was designed with a radically different, open, and distributed technological architecture (Clark, 2018; Leiner et al., 1997; van Schewick, 2010). This required a comparable networked governance structure that was sufficiently flexible to evolve and adapt over time (Brousseau, Marzouki, & Méadel, 2012; DeNardis, 2009, 2014; Mueller, 2002, 2010).

Prior technological transformations, such as the steam engine, electricity, and computing, respectively, were all influenced by public policies. Awareness of the importance of good governance increased in the late twentieth century and was progressively recognized as a prerequisite for well-functioning markets (Ostrom, 2010; Roth, 2018). OECD countries started to share information about good regulation and other international bodies such as the International Telecommunication Union (ITU) and World Bank also emphasized its importance (e.g., Odugbemi & Jacobson, 2008; OECD, 1995; OECD, 2012). During the past decade, a consensus emerged among national and international policymakers that the current, interrelated breakthroughs across digital, biological, and physical spheres might be impeded unless additional regulatory reforms were adopted. Approaches, such as adaptive regulation, anticipatory regulation, outcome-based regulation, experimental regulation, and data-driven regulation are the latest responses to these challenges (World Economic Forum, 2020).

Today, individual nations and the global community consider innovation as a critical component of solutions to grand challenges. Hope is widespread that innovation can mitigate climate change, youth unemployment, food security, problems of aging populations, and immigration.

The Coronavirus pandemic has highlighted the importance of flexible adaptation to dynamically evolving crises and also revealed the failures associated with applying old policies to new problems (De Waal, 2021). Even though these recent experiences have added urgency to the interest in agile regulation, the discussion has earlier origins. It reflects general concerns that the Fourth Industrial Revolution will only flourish in an appropriate institutional framework (Schwab, 2016). Specifically, the present scientific and technological breakthroughs follow a different, more ubiquitous, and cross-sectoral pattern than earlier periods of technological change. This requires novel, governance responses.

It is known that traditional regulation fails to achieve its purposes under certain conditions. Information asymmetries, bureaucratic processes, legal disputes, rent-seeking, and regulatory capture can contribute to considerable regulatory inertia and rigidity (Dixit, 1996). Such forms of government failure are potentially aggravated in the digital economy and the global, socio-technical systems that propel it. Efforts to address these weaknesses of government regulation and public ownership can be found throughout history, but especially since the 1970s. In the 1980s, performance-based and other forms of incentive regulation, such as price caps, were introduced (Sappington & Weisman, 2010). Researchers and practitioners developed new frameworks, such as responsive regulation (Ayres & Braithwaite, 1992), smart regulation (Gunningham, Grabosky, & Sinclair, 1998), experimentalist governance (Sabel & Zeitlin, 2010), and several other proposals (Brass, 2021). Agile regulation builds on these earlier institutional innovations and experiments.

New forms of regulation are not only a response to changing patterns of technological change, but they are also a reaction to increasing interdependencies among players and stakeholders that require novel forms of coordination and governance. For example, one nation alone cannot fully address the security concerns related to the Internet of Things (IoT). Those concerns would benefit from collaboration to promulgate improved international standards. Similarly, globally mobile companies create growing transnational interdependencies that reduce the ability of one individual nation or player to influence outcomes. This increases uncertainty about the effects of governance. Traditional efforts to control the system, typically designed using static optimization approaches, decline in effectiveness under these conditions. Instead, governance increasingly resembles a tuning device that continuously adjusts the course of action or the form of intervention to experience. The agile regulation discussion is also a response to this changing dynamic nature of governance.

3. Realigning governance and techno-economic conditions

Institutional analyses of governance suggest that the of alignment between regulation and the technological and economic conditions of a sector is critical. Proper alignment may even be more important for performance than the specific instruments that are being deployed (Finger, Groenewegen, & Künneke, 2005; Künneke, Groenewegen, & Ménard, 2010). Alignment requires that the components of the institutional matrix at the level of individual agents, organizations, sectoral regulation and legislation, and overarching institutional arrangements are in a workable relationship with each other (e.g., Finger et al., 2005; Künneke et al., 2010; Rodrik, 2007). In dynamic systems, several sets of governance constellations typically exist that can solve this institutional design problem. However, different workable governance constellations will correspond to characteristic performance patterns and trade-offs. For example, a laissez faire governance approach may result in a high rate of entrepreneurial innovation and high inequality of its benefits. In contrast, an alternative option characterized by stronger regulatory interventions might be associated with slower technological change and higher equality.

Agile methods seek to improve alignment between the economic and technological conditions of the Fourth Industrial Revolution and its governance. One common characteristic of agile regulation is the willingness of governments and other stakeholders to experiment with new regulatory modes and methods, such as outcome-based approaches. A second shared trait is the periodic, systematic review of existing regulations to eliminate obsolete and superfluous interventions and design instruments that are better suited to emerging industry conditions. Finally, agile methods of regulation rely on the systematic monitoring and analysis of the effects of interventions to adjust governance to industry conditions and the economic and social objectives of regulation (OECD, 2021b).

Whereas these aspects of adaptive regulation are reactive, *anticipatory* regulation is envisioned as a proactive, forward-looking approach. In this method, horizon-scanning approaches are used to identify emerging technologies and services. In combination with scenario-planning technologies, the potential economic and social value of an innovation can be assessed. Examination of potential regulatory roadblocks that could be removed or of potential risks and harms that might require new safeguards helps to inform policy adaptation. These steps are repeated to facilitate the adaptation of governance in ways that allow the realization of the benefits of innovations while mitigating their potential harms. Several nations have established capabilities to assess systematically the policy challenges and

developments that are related to emerging technologies (e.g., the Swedish Committee for Technological Innovation and Ethics) (World Economic Forum, 2020, pp. 9–13).

Outcome-based regulation seeks to identify desirable outcomes without prescribing technical, economic, and business models to achieve them. Where appropriate, such goal-based approaches have the advantage to be technologically neutral and to release entrepreneurial energy in pursuit of the goals. However, in a dynamically changing environment, setting and monitoring appropriate goals may be challenging if not impossible. Whereas ambitious goals can drive innovation, overly ambitious goals may produce repeated failure. Too cautious goals, on the other hand, may define low performance benchmarks that become focal points for players and reduce their incentives to outperform them. Some of these problems may be addressed by allowing deviations from regulations for limited periods or by introducing sunset clauses. Finally, in areas with high uncertainty, government can adopt soft law tools, such as regulatory guidance, to reduce uncertainty (e.g., AI ethics guidelines).

Experimental regulation refers to approaches that are deliberately designed to support and enable innovative products and services. Examples include the facilitation of testing under regulatory supervision in test beds, laboratories, and regulatory “sandboxes.” Another approach is the establishment of innovation hubs that can help with the launching of innovations. Finally, several countries are experimenting with one-stop shops or single points of contact for entrepreneurs that simplify the complexity of navigating rules and regulations that govern startups. There is a potential tension between the legitimate desire of businesses for stable regulation and the benefits of experimentation. Where experimental regulation leads to subsequent regulatory reform, it is necessary to ensure that related regulations are adapted as well. This may apply at the national level, but in many cases, it may also require international coordination.

A portfolio of additional methods to address the alignment problem is in discussion. *Data-driven* regulation seeks to link interventions closely to empirical evidence. The resulting, reactive responses may be appropriate for certain scenarios. However, if a policy seeks to change the future course of a system, big data analytical methods may be limiting, because they inevitably reflect the past and the status quo. *Algorithmic* regulation goes a step further and would put certain types of interventions on autopilot. Regulatory tasks would be delegated to algorithms with some level of decision-making capability. Again, this approach will work in certain limited scenarios. However, it faces all the risks known from reliance on algorithms and discussions about emerging forms of artificial intelligence.

For example, racial, educational, or income biases inherent in the design and training of algorithms may be replicated or amplified (Eubanks, 2017; Latzer & Just, 2020).

Arguments for agility are often based on pragmatic proposals. However, several aspects of the alignment problem are only incompletely addressed. It is not always evident whether a specific method of flexible regulation is appropriate (e.g., incentive compatible) to address the problem at hand. More conceptual and empirical work is needed to examine whether specific tools are effective and efficient. Moreover, agile regulation not only has benefits, but it also imposes costs on regulators and the affected stakeholders. The rational design of agile regulatory tools must assure that the benefits outweigh the costs, including any transaction and adaptation costs associated with frequently changing regulations. We will discuss these challenges using the example of the information and communication sector.

4. Agile regulation in the digital economy

The concept of agile regulation developed as a response to rapid, intersectoral, scientific change and innovation. However, there are parallels and similarities between the Fourth Industrial Revolution and the governance challenges of the digital economy. Like the Fourth Industrial Revolution, the digital economy is shaped by fast-paced, accelerating technological change. Digital technology is highly plastic, which has tremendously expanded the innovation opportunities in the digital economy. Advanced forms of digital value generation, for example in manufacturing, agriculture, and advanced mobility, rely on services that are configured via software-defined, virtual networks. These typically integrate functions from interrelated layers of the innovation ecosystems, including the physical network, logical functions such as routing, and applications and services. Often, effective integration requires coordination across a large number of complementary players. However, the social and economic consequences of these innovations are often not known in advance or are only partially understood. In these conditions, agile regulation is better suited than traditional regulation.

Historically, the introduction of an adaptive governance model was a major policy innovation that contributed to the success and eventual rapid growth of the Internet (Garcia, 2016). As the telecommunications sector evolved from an era of monopoly to a competitive organization, traditional forms of command-and-control regulation that aimed at the

maximization of static, economic efficiency were increasingly inadequate. In the new environment, dynamic change and innovation were major drivers of efficiency gains, but they were only implicitly considered in the design of traditional regulation (Bauer, 2019).

With the increasing awareness of the importance of innovation, regulation was gradually reoriented from reliance on traditional control theory and engineering to principles of adaptive governance. Many of the design features of agile regulation were articulated in early experiments with adaptive governance and realized in the institutions and mechanisms of Internet governance. Among them were a stronger role for bottom-up governance processes, a reliance on the experience and insights from peer groups of experts (e.g., in the Internet Engineering Task Force—IETF), and a more reflexive approach that combined continuous monitoring of outcomes with adaptations of governance (e.g., Cherry & Bauer, 2004; Mueller, 2010).

New governance challenges arose as the Internet matured and all-Internet Protocol (all-IP) network architectures became the ubiquitous basis of digital communications. Market dominance and the abuse of market power have become pressing issues of digital economy policy. Search, cloud computing, e-commerce, and entertainment markets have all become highly concentrated. In response, several countries are envisioning hybrid arrangements that combine elements of traditional regulation with elements of competition policy. Hybrid arrangements hope to overcome the dual challenges of traditional approaches to regulation and competition policy: The former is perceived as inappropriate for technologically dynamic industries and the latter is seen as too slow and ineffective. New, institutional solutions seek to cover a middle ground. For example, anticipatory competition policy, as encoded in the European Digital Markets Act, uses behavioral constraints on dominant players to address market power (Geradin, 2021). Ex post regulation, more compatible with the U.S. legal system, establishes a framework for market outcomes but allows interventions only when actual outcomes deviate in unacceptable ways (Yoo, 2012).

During the past decade, expanding broadband access to remote regions and low-income populations arose as a new governance problem. Only a few countries updated universal service funding mechanisms, originally introduced in the twentieth century to connect households to telephone service, to include broadband (Prado & Bauer, 2021). In many cases, spectrum auctions are designed to include public interest obligations for mobile broadband, such as roll out and coverage requirements (Cave & Nicholls, 2017). Because these obligations can be flexibly designed and

modified in subsequent auctions, they may be considered a variant of flexible regulation. An increasing number of stakeholders do not envision mobile broadband as a viable substitute for fixed broadband connections. In countries that have not adopted national policies to support universal fixed broadband access, lower levels of government, including cities and municipalities, are increasingly stepping in to assume responsibility.

For example, in the United States, several states and an increasing number of local communities have become highly active in fixed broadband policy. This is, first, a response to policy deficiencies and failures at the federal level, such as the reliance of many subsidy programs, such as the Connect America Fund (CAF) or the Rural Digital Opportunities Fund (RDOF) on erroneous broadband mapping data (Ali, 2021). Second, the sprouting of local initiatives is a response to the promise of massive amounts of infrastructure funding from the federal government. One might be tempted to see the American approach as a dysfunctional patchwork of initiatives. However, it could also be viewed as a polycentric approach or a laboratory in which local, natural experiments are conducted. The varying local conditions may facilitate the development of workable best practices that can be replicated elsewhere (Teske, 1990; Whitacre & Gallardo, 2020).

Emerging technologies, such as fifth and sixth generation (5G, 6G) wireless, the Internet of Things (IoT), and the advanced cyber-physical systems of the Internet of Intelligence (IoI) involve coordination between a large number of complementary players (Bauer & Bohlin, 2022). Initial stages of innovation require experimentation to find the organizational and institutional arrangements that can effectively achieve the required coordination. When such workable arrangements have been found, it may be possible to achieve coordination via standards, modular technology architectures, and market-driven contacts. However, a diversity of experiments and flexibility are important during the initial stages. An increasing number of countries have allowed a differentiation of regulatory rules to accommodate such experiments. An example is the diversification of 5G spectrum policy across Europe, often intended to facilitate local experiments with innovative applications and use cases (Knieps & Bauer, 2022; Matinmikko-Blue, Yrjölä, Ahokangas, & Hämmäinen, 2021). Thus, flexibility and agility support innovation experiments, especially during the early development of complex technologies.

Another example that illustrates the potential usefulness of agile regulation in the digital economy is the governance of artificial intelligence (AI). Currently available, weak AI is a cheap prediction and decision support technology (Agrawal, Gans, & Goldfarb, 2018). Its rapidly declining cost

and the ability to build algorithms with only limited programming knowledge have contributed to rapid adoptions across many uses. Although there are strong movements toward open and ethical AI, most algorithms are proprietary and, from a user's perspective, constitute a black box. To govern the application of AI, agile regulation would be superior to traditional forms of oversight. Methods such as sandboxing to experiment with new algorithms in a limited fashion would allow innovation while establishing safeguards against abuses. This would also allow the design of standards and governance mechanisms that protect important public interest goals. Traditional forms of regulation could be limited to absolutely critical issues, such as constraining uses that have a high likelihood to cause more harm than good.

5. Toward stronger theoretical foundations and practices

Much of the discussion about agile regulation focuses on scientific progress and how its benefits can best be harnessed. Given the early stage of institutional experiments with agile methods of regulation, it is not surprising that many aspects need additional work. This section addresses a few key issues that would benefit from more discussion: the reconciliation of potentially contradictory roles of regulation as a restraining tool and as a supportive tool; the challenges associated with understanding the effects of agile regulation on sector outcomes; the design of appropriate instruments that address the trade-offs between agility and the increased transaction costs of regulation; and the need to reflect more critically on the rate and direction of innovation. All of these issues would also benefit from stronger theoretical foundations.

The design of agile regulation is a special case of the more general institutional design problem mentioned above: the alignment of governance arrangements with the technical, economic, and social conditions of a socio-technical system. Recent advances in economics highlight the importance of the design of appropriate rules for markets (Maskin, 2008; Milgrom, 2017; Ostrom, 2010; Roth, 2018). Current agile regulation proposals do not build on this body of conceptual knowledge, although they may point to it in a metaphorical way. The present discussion is driven by practitioners and the willingness to experiment with new approaches to governance. It would be desirable to link the discussion and the design of governance to what is known about the desirable and undesirable effects of economic mechanisms and incentives. Not only can this help to avoid some of the known problems (Bowles, 2016; Bowles & Polania-Reyes,

2012), but it can also help inform the choice of the most efficient, agile governance arrangement for a sector.

A second challenge is that the dual goal to support innovation while mitigating the potential harms of emerging technologies often requires a mix of restraining and promotional forms of governance. The ongoing policy discussions recognize this. Yet, the diversity of scenarios in which regulation may be invoked and the specific form of tensions and trade-offs between restraining and promoting roles remain underexplored. This will typically require going beyond a single instrument approach to employ a combination of *ex ante*, *ex post*, and agile instruments (Gunningham & Sinclair, 1999). The combination of instruments will depend on the policy challenge at hand. There will typically be more than one feasible option and selecting the most efficient solution may not be straightforward. For example, concerns about the ability of digital platforms to manipulate the market environment of competitors could be addressed by forms of *ex ante* regulation, by anticipatory (*ex ante*) competition policy, by methods of *ex post* regulation, or by traditional forms of (*ex post*) competition enforcement. Each of these choices will affect the competitive and innovation dynamics of the sector differently (Prado, 2022).

This points to the related problem of developing a robust understanding of the effects of different governance approaches on outcomes. In dynamic industries with interdependent, complementary innovation processes, this will typically require an understanding of the direction and strengths of multiple first- and second-round effects (Bauer & Bohlin, 2022). *Ex ante* regulation is contingent on finding an answer to this dynamic governance problem before an appropriate intervention can be designed. No widely accepted methods and tools exist currently to overcome the associated tremendous information and modeling challenges. Consequently, agile forms of regulation are often applied on a limited spatial and temporal scale to gain practical experience with their effects. Moreover, they explicitly design feedback loops to monitor outcomes with the goal to adjust the strength and form of intervention. Hence, they are particularly well-suited to govern processes in dynamic systems with high levels of interdependence and complementarity. They have their own limits when issues arise that cannot easily be addressed experimentally and locally. For example, it may be difficult to sandbox transnational companies and to confine them to local operations.

The transaction costs associated with agile regulation are another aspect that is insufficiently addressed in the current debate. In many cases, agile regulation reduces regulatory burdens, at least during an experimental, learning period. At some point, however, governance choices will have to

be applied more broadly at a sector or even trans-sectoral level. At this point, a trade-off becomes more pertinent: that changes in the regulatory model will affect the transaction costs of a growing number of players in the sector. After an initial setup of the market, further changes in regulations will typically require adaptations of business processes, marketing methods, pricing, etc. They will also affect the expected profitability of investment and innovation projects. Not all players will be affected equally. Those who need to commit resources for the long-term will prefer stable regulation that does not change frequently; they may seek to capture regulation in their own favor. Entrepreneurs seeking to enter a market may support more frequent changes of regulations. A migration to and continued reliance on agile forms of regulation will be meaningful only if the benefits outweigh the costs of a transition, including transaction costs. This implies that the optimal rate of updating – the optimal agility – is contingent on the associated costs and benefits. Faster and more agile regulation is not always better. Sub-optimally high levels of agility may even contribute to a fragmentation of the innovation ecosystem (e.g., Stocker, Knieps, & Dietzel, 2021).

A strong belief in the benefits of innovation energizes much of the enthusiasm for agile regulation. Rarely is this fundamental trust in the desirability of innovation questioned, even if innovation is not seen as an end in itself, but as a means to achieve ends such as sustainable growth, well-being, and to address global challenges such as climate change (OECD, 2021b). However, not all innovation promotes the public interest. Private sector, commercial innovation may or may not be in the subset that is also in the public interest (McGuinness & Schank, 2021; Toyama, 2015). The discussion about social media platforms that exploit insights from behavioral social science to manipulate users shows that some profit-increasing innovations (in this case, process and business model innovations) may be afflicted with negative, societal externalities. Likewise, competition in markets with first-mover advantages often leads to the release of software and other digital products whose security has not been appropriately tested. Such products can subsequently cause potentially very damaging cybersecurity vulnerabilities (Bauer & Van Eeten, 2009).

Digital innovation is highly malleable and often spawns technologies that have multiple beneficial and potentially harmful uses. For example, social media connects individuals and groups, but they are also instruments of surveillance (Zuboff, 2019). There is a risk that in the enthusiasm for the new and for agile regulation, an innovation trajectory is pursued that is biased in favor of the interests of for-profit stakeholders. Here, public policy faces the challenge to guide agile entrepreneurship in direc-

tions that are congruent with the public interest. One way to achieve this might be to channel and deliberately slow innovation (Toyama, 2015). There are costs associated with this approach, as it reduces the diversity of innovation paths that are explored. Thus, potential benefits and costs will have to be carefully examined. Another option might be to facilitate institutional diversity, which creates players with different motivations and incentives. Cooperatives or B-corporations that work in parallel with for-profit companies will choose a different innovation path, generating knowledge about what else could be. Similarly, publicly funded projects may put pressure on commercial innovators to explore alternative innovation trajectories (Mazzucato, 2013, 2021). Agile regulation embedded in such a larger framework of deliberate institutional diversity, and designed as an instrument of innovation policy, seems to have a higher chance of generating a balanced innovation path.

6. Conclusion

Interest in agile regulation is related to the hope that scientific progress and innovation at the heart of the Fourth Industrial Revolution will help overcome the grand challenges faced by individual nations and the global community. This chapter has critically reviewed these claims and explored the potential role of agile and other forms of flexible regulation in the information and communications sector, which is one of the drivers of the Fourth Industrial Revolution.

Adaptive regulation was one of the innovations that helped propel the early Internet. Recent experience reveals several additional areas, such as experiments with localized, industrial applications of 5G that illustrate the potential usefulness of flexible forms of governance. Flexibility is part of many discussions of the major, governance challenges, including the taming of market power of digital platforms, the universalization of broadband access, the promotion of innovation, and the mitigation of potentially harmful effects of computer-mediated communications and social media. Overall, however, the complexity and multifaceted nature of advanced information and communication technologies suggest that single-instrument approaches to policy have limited applicability. Rather, governance arrangements that combine multiple tools are needed.

There is room and use for agile approaches to regulation in the digital economy, but they are embedded in and will complement other forms of governance. The most important, overarching insight is that all forms of governance need to be approached as dynamic learning systems, in

which a process of monitoring is combined with arrangements that facilitate systematic learning. In such a framework, feedback is used to adapt governance in response to system outcomes. This transformation is only at the beginning, and much can be learned from practical experience with agile regulation and its effects on outcomes. Practitioners drive the current debate, but considerable room exists for collaboration with researchers to strengthen the conceptual foundations of agile regulation and to refine the implementation of practical tools.

References

- Agrawal, A., Gans, J. S., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Boston, MA: Harvard Business Review Press.
- Ali, C. (2021). *Farm fresh broadband: The politics of rural connectivity*. Cambridge, MA: MIT Press.
- Ayres, I., & Braithwaite, J. (1992). *Responsive regulation: Transcending the deregulation debate*. New York: Oxford University Press.
- Bauer, J. M. (2019). Regulation and digital innovation. In G. Knieps & V. Stocker (Eds.), *The future of the Internet-Innovation, Integration and Sustainability* (pp. 58–88). Baden-Baden: Nomos.
- Bauer, J. M., & Bohlin, E. (2022). Regulation and innovation in 5G markets. *Telecommunications Policy*, 46(4), 102260. <https://doi.org/10.1016/j.telpol.2021.102260>.
- Bauer, J. M., & Van Eeten, M., J. G. (2009). Cybersecurity: Stakeholder incentives, externalities, and policy options. *Telecommunications Policy*, 33(10/11), 706–719. <https://doi.org/10.1016/j.telpol.2009.09.001>.
- Bowles, S. (2016). *The moral economy: Why good incentives are not substitute for good citizens*. New Haven, CT: Yale University Press.
- Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, 50(2), 368–425. <https://doi.org/10.1257/jel.36850.2>.
- Brass, I. (2021, July 26). Agile, Anticipatory, Adaptive: Regulatory Responses to Technology Disruptions and Scientific Innovations [Webinar]. Theories of Regulation & Governance. <https://www.youtube.com/watch?v=acwgYzASnK4>.
- Brass, I., & Sowell, J. H. (2020). Adaptive governance for the Internet of Things: Coping with emerging security risks. *Regulation & Governance*, 15(4), 1092–1110. <https://doi.org/10.1111/rego.12343>.
- Brousseau, E., Marzouki, M., & Méadel, C. (Eds.). (2012). *Governance, regulation, and powers on the Internet*. Cambridge, UK: Cambridge University Press.
- Brunner, R. D. (2010). Adaptive governance as a reform strategy. *Policy Sciences*, 43(4), 301–341. <https://doi.org/10.1007/s11077-010-9117-z>.

- Cave, M., & Nicholls, R. (2017). The use of spectrum auctions to attain multiple objectives: Policy implications. *Telecommunications Policy*, 41(5–6), 367–378. <https://doi.org/10.1016/j.telpol.2016.12.010>.
- Cherry, B. A., & Bauer, J. M. (2004). *Adaptive regulation: contours of a policy model for the Internet economy*. Quello Center Working Paper 04–05, East Lansing, MI: Michigan State University.
- Clark, D. D. (2018). *Designing an Internet*. Cambridge, MA: MIT Press.
- De Waal, A. (2021). *New pandemics, old politics: Two hundred years of war on disease and its alternatives*. Cambridge, UK: Polity.
- DeNardis, L. (2009). *Protocol politics: The globalization of Internet governance*. Cambridge, MA: MIT Press.
- DeNardis, L. (2014). *The global war for Internet governance*. New Haven, CT: Yale University Press.
- Dixit, A. K. (1996). *The making of economic policy: A transaction cost politics perspective*. Cambridge, MA: MIT Press.
- Eubanks, V. (2017). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press.
- Finger, M., Groenewegen, J., & Künneke, R. (2005). The quest for coherence between institutions and technologies in infrastructures. *Journal of Network Industries*, 6(4), 227–259. <https://doi.org/10.1177/178359170500600402>.
- Folke, C., Hahn, T., Olsson, P., & Norberg, J. (2005). Adaptive governance of social-ecological systems. *Annual Review of Environment and Resources*, 30, 441–473. <https://doi.org/10.1146/annurev.energy.30.050504.144511>.
- Garcia, D. L. (2016). The evolution of the Internet: A socioeconomic account. In J. M. Bauer & M. Latzer (Eds.), *Handbook on the economics of the Internet* (pp. 529–552). Cheltenham, UK; Northampton, MA: Edward Elgar.
- Geradin, D. (2021). *What Is a Digital Gatekeeper? Which Platforms Should Be Captured by the EC Proposal For A Digital Market Act?* (SSRN Scholarly Paper ID 3788152). Social Science Research Network. <https://ssrn.com/abstract=3788152>.
- Gunningham, N., Grabosky, P., & Sinclair, D. (1998). *Smart regulation*. Oxford: Oxford University Press.
- Gunningham, N., & Sinclair, D. (1999). Regulatory pluralism: Designing policy mixes for environmental protection. *Law & Policy*, 21(1), 49–76. <https://doi.org/10.1111/1467-9930.00065>.
- Knieps, G., & Bauer, J. M. (2022). Internet of things and the economics of 5G-based local industrial networks. *Telecommunications Policy*, 46(4), 102261. <https://doi.org/10.1016/j.telpol.2021.102261>.
- Künneke, R., Groenewegen, J., & Ménard, C. (2010). Aligning modes of organization with technology: Critical transactions in the reform of infrastructures. *Journal of Economic Behavior & Organization*, 75(3), 494–505. <https://doi.org/10.1016/j.jebo.2010.05.009>.

- Latzer, M., & Just, N. (2020). Governance by and of algorithms on the Internet: Impact and consequences. In J. F. Nussbaum (Ed.), *Oxford Research Encyclopedia Communication*: Oxford University Press. <https://doi.org/10.1093/acrefore/9780190228613.013.904>.
- Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., Wolff, S. S. (1997). The past and future history of the Internet. *Communications of the ACM*, 40(2), 102–108. <https://doi.org/10.1145/253671.253741>.
- Majone, G. (1996). The rise of statutory regulation in Europe. In G. Majone (Ed.), *Regulating Europe* (pp. 47–60). London: Routledge.
- Maskin, E. S. (2008). Mechanism design: How to implement social goals. *American Economic Review*, 98(3), 567–576. <https://www.jstor.org/stable/29730086>.
- Matinmikko-Blue, M., Yrjölä, S., Ahokangas, P., & Hämmäinen, H. (2021). *Analysis of 5G spectrum awarding decisions: How do different countries consider emerging local 5G networks?* 2021 ITS Biennial Conference, Gothenburg.
- Mazzucato, M. (2013). *The entrepreneurial state: Debunking public vs. private sector myth*. London: Anthem Press.
- Mazzucato, M. (2021). *Mission economy: A moonshot guide to changing capitalism*. New York: Harper Business.
- McGuinness, T. D., & Schank, H. (2021). *Power to the public: The promise of public interest technology*. Princeton, NJ: Princeton University Press.
- Milgrom, P. (2017). *Discovering prices: Auction design in markets with complex constraints*. New York: Columbia University Press.
- Mueller, M. L. (2002). *Ruling the root: Internet governance and the taming of cyberspace*. Cambridge, MA: MIT Press.
- Mueller, M. L. (2010). *Networks and states: The global politics of Internet governance*. Cambridge, MA: MIT Press.
- Odugbemi, S., & Jacobson, T. (2008). Governance reform under real-world conditions: Citizens, stakeholders, and voice. Washington, DC: World Bank. <https://opendata.worldbank.org/handle/10986/6513>.
- OECD (1995). Recommendation of the Council on improving the quality of government regulation (OECD Legal Instruments). Organisation for Economic Co-operation and Development. <https://legalinstruments.oecd.org/public/doc/128/128.en.pdf>.
- OECD (2012). Recommendation of the Council on regulatory policy and governance. Paris: OECD Publishing. <http://dx.doi.org/10.1787/9789264209022-en>.
- OECD (2021a). *OECD regulatory policy outlook 2021*. Paris: OECD Publishing. <https://doi.org/10.1787/38b0fdb1-en>.
- OECD (2021b). Recommendation of the Council for agile regulatory governance to harness innovation. OECD/LEGAL/0464. <https://legalinstruments.oecd.org/api/print?id=669&lang=en>.

- Ostrom, E. (2010). Beyond markets and states: Polycentric governance of complex economic systems. *American Economic Review*, 100(3), 641–672. <https://www.jstor.org/stable/27871226>.
- Prado, T. S. (2022). Safeguarding Competition in Digital Markets: A Comparative Analysis of Emerging Policy and Regulatory Regimes. Quello Center Working Paper No. 05, 2022. <https://ssrn.com/abstract=4137588>.
- Prado, T. S., & Bauer, J. M. (2021). Improving broadband policy design using market data: A general framework and an application to Brazil. *Telecommunications Policy*, 45(4), 102111. <https://doi.org/10.1016/j.telpol.2021.102111>.
- Rodrik, D. (2007). *One economics many recipes: Globalization, institutions, and economic growth*. Princeton, NJ: Princeton University Press.
- Roth, A. E. (2018). Marketplaces, markets, and market design. *American Economic Review*, 108(7), 1609–1658. <https://doi.org/10.1257/aer.108.7.1609>.
- Sabel, C. F., & Zeitlin, J. (Eds.). (2010). *Experimentalist governance in the European Union: Towards a new architecture*. Oxford: Oxford University Press.
- Sappington, D. E. M., & Weisman, D. L. (2010). Price cap regulation: What have we learned from 25 years of experience in the telecommunications industry? *Journal of Regulatory Economics*, 38(3), 227–257. <https://doi.org/10.1007/s11149-010-9133-0>.
- Schwab, K. (2016). *The fourth industrial revolution*. Geneva: World Economic Forum.
- Steelman, T. A. (2016). Adaptive governance. In C. Ansell & J. Torfing (Eds.), *Handbook on theories of governance* (pp. 538–550). Cheltenham, UK & Northampton, MA: Edward Elgar.
- Stocker, V., Knieps, G., & Dietzel, C. (2021). The Rise and Evolution of Clouds and Private Networks–Internet Interconnection, Ecosystem Fragmentation. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn3910108>.
- Teske, P. E. (1990). *After divestiture: The political economy of state telecommunications regulation*. Albany, N.Y.: State University of New York Press.
- Toyama, K. (2015). *Geek heresy: Rescuing social change from the cult of technology*. New York: PublicAffairs.
- van Schewick, B. (2010). *Internet architecture and innovation*. Cambridge, MA: MIT Press.
- Viscusi, W. K., Harrington, J. E., Jr., & Sappington, D. E. M. (2018). *Economics of regulation and antitrust*. Cambridge, MA: MIT Press.
- Whitacre, B., & Gallardo, R. (2020). State broadband policy: Impacts on availability. *Telecommunications Policy*, 44(9), 102025. <https://doi.org/10.1016/j.telpol.2020.102025>.
- World Economic Forum. (2020). *Agile regulation for the Fourth Industrial Revolution: A toolkit for regulators*. <https://www.weforum.org/about/agile-regulation-for-the-fourth-industrial-revolution-a-toolkit-for-regulators>.
- Yoo, C. S. (2012). *The dynamic Internet: How technology, users, and businesses are transforming the network*. Washington, D.C.: AEI Press.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York, NY: Random House.

Preisbildung bei der Zusammenschaltung im Internet

*Falk von Bornstaedt*¹

Abstract

Zusammenschaltungen sind das Kernelement des Internets. Die umfassendste und kostenpflichtige Variante *IP-Transit* stellt die vollständige globale Erreichbarkeit sicher, meist mit Qualitätsgarantien. Das *Peering* erlaubt den kostenlosen Austausch zwischen zwei Netzen, überwiegend ohne Qualitätsgarantien. Bei zunehmender Anzahl von Applikationen mit Qualitätsanspruch und stark gesunkenen Kosten für Transit sollte das Dogma der „kostenlosen Zusammenschaltung“ überprüft werden.

Das Internet hat sich seit 2010 erheblich verändert hin zu einer Video-Distributionsplattform. Die Seite der Inhaltenanbieter hat sich immer mehr auf wenige Unternehmen konzentriert, denen stehen in Europa zahlreiche Endkundenprovider gegenüber. Es ist zu untersuchen, ob sich in der Folge auch die kommerziellen Rahmenbedingungen fortentwickeln müssen.

Der Verzicht auf das Abrechnen der 5 % Zeitschlitz (time slots) mit der höchsten Auslastung bei Abrechnung nach der 95-Perzentil-Regel gibt ökonomisch den falschen Anreiz, zusätzliche Verkehre ausgerechnet in die Zeiten der Spitzenlast zu legen. Das ist ökonomisch und ökologisch ein Problem. Es sollte ein innovatives Preissystem entwickelt werden, das die falschen Anreize vermeidet.

Keywords:

IP-Transit, Peering, Zusammenschaltung, Interconnection

1. Einführung

Das Internet ist ein Netz, bestehend aus ca. 74000 Autonomen Systemen (AS) (vgl. CIDR, n.d.; hier werden nur die tatsächlich aktiven Netze gezählt). Diese müssen alle untereinander verbunden sein, um die Kundenanforderung der umfassenden globalen Erreichbarkeit sicherzustellen. Es

1 BENOCS; fbornstaedt@benocs.com.

ist nicht praktikabel, jedes Netz mit jedem Netz direkt zu verbinden, dies würde bedeuten, $0,5n(n-1)$ Verbindungen zu schaffen, das sind etwa 2,7 Milliarden Zusammenschaltungen. Es muss eine Konzentration stattfinden, wo Netze Verkehre auch für andere Netze durchleiten.

In den Anfängen des Internet gewährten die Netze einander umfassende kostenlose Zusammenschaltung (=Interconnection). Mit der Kommerzialisierung des Internets endete dieses Regime. Die großen Provider, die ein umfangreiches Netz finanzierten, wollten dieses nur gegen Entgelt Dritten zur Mitbenutzung überlassen. Die großen US-Carrier entdeckten zuerst, dass sich das Internet monetarisieren ließ.

Zusammenschaltungen verschiedener eigenständiger Netze erfolgen auf Basis des Border Gateway Protocol Version 4 (BGP4). Jedes Netz teilt dem Partner bei der Zusammenschaltung mit („annonciert“), über welche Routen es verfügt. Dabei werden die vom Partner annoncierten Routen in die eigene Routingtabelle übernommen. Dieser Austausch benötigt Vertrauen, falsch annoncierte Routen können dramatische Auswirkungen haben.

Tabelle 1: Formen der Zusammenschaltung

Routen annonciert	Zusammenschaltung	Bepreisung
Nur die Routen des eigenen AS und der eigenen Kunden	Peering	Überwiegend „settlement free“
Alle Routen einer bestimmten Region	Limited IP-Transit	Reduzierter IP-Transit Preis
Volle globale Routingtabelle	IP-Transit	IP-Transit Preis

Quelle: Eigene Darstellung.

Um die Komplexität zu reduzieren, beschränkt man sich im Wesentlichen auf Peering und Transit. Beim gegenwärtigen Stand der IT Tools könnte ein komplizierteres System dazu führen, dass bestimmte Adressbereiche nicht mehr von überall erreichbar sind. Auch im Störfall ist es sehr hilfreich, überschaubare Strukturen zu haben und nicht viele Sonderlösungen.

Die kostenpflichtige Zusammenschaltung im Internet könnte man als homogenes Massengut (commodity) ansehen, soweit man Qualitätsunterschiede ignoriert. Die Markttransparenz bezüglich der Qualität hat sich stetig verbessert, durch die Dynamik des Internets sind Qualitätsaussagen

schwer zu generalisieren. Je nach Branche und Anwendung gelten stark unterschiedliche Anforderungen. Bei einer Videokonferenz kommt es darauf an, einen stetigen Fluss der Pakete zu gewährleisten, verlorene Pakete ein paar Sekunden später erneut zu übertragen, macht keinen Sinn. Bei Onlineeinkäufen hingegen ist es wichtig, dass alle Pakete ankommen, und sei es ein paar Sekunden später.

Das generelle Preisniveau für die Zusammenschaltung insgesamt sinkt sowohl durch einen höheren Anteil der kostenlosen Zusammenschaltung als auch durch sinkende Preise bei IP-Transit und bezahltem Peering. Das Preisniveau 2015 betrug nur noch ein Tausendstel der 1200 \$/Mbps/Monat von 1998. Der Preisverfall wird getrieben durch den technischen Fortschritt und einen starken Wettbewerb der Transitanbieter. Regionale Unterschiede verringern sich zunehmend, sie sind aber immer noch relevant, so kostet z.B. ein 10 Gbps Port in Singapur 3,8-mal so viel wie in London (TeleGeography, 2020).

Nach Angaben von Cloudflare (Prince & Rao, 2021) sind in den letzten 10 Jahren die IP-Transit-Preise um 23 % pro Jahr gefallen. Die Großhandelspreise für Bandbreiten sind in den 10 Jahren insgesamt um 93 % gefallen, die Preise für das Ausspielen von Amazon Web Services (AWS) Inhalten aber nur um 25 %.

Die Internet Exchange Points (IXPs) haben ebenfalls stark zum Preisverfall beigetragen, weil sie den Wettbewerb befördert haben. Das gilt insbesondere für die drei großen europäischen Austauschpunkte in Frankfurt, Amsterdam und London. Es gibt weltweit ca. 1000 Internet-Austauschpunkte.

Ein 10 Gbps Anschluss mit einem Jahresvertrag am AMS-IX kostet monatlich 650€ ohne Steuern und ohne die Inhaus-Verkabelung. Rechnet man mit einem 95 Perzentil Wert der Nutzung von 6,5 Gbps, so ergeben sich Kosten von 0,10€ pro Mbps pro Monat. Das ist weniger als die Hälfte der Kosten eines durchschnittlichen Transitports in Frankfurt (Coll, 2021).

Grundsätzlich ist es effizient, ein Produkt an den Kosten orientiert zu bepreisen. Dann geben die Preise die richtigen Signale an alle Marktteilnehmer. In den Märkten der Internet-Zusammenschaltung gibt es positive Netzwerkexternalitäten. Bei IP-Transit erhöht jeder neu angeschaltete Kunde den Nutzen der bisherigen Kunden. Das verführt dazu, Wachstum zu priorisieren und das Netz um nahezu jeden Preis zu füllen. Die Grenzkosten eines zusätzlichen Kunden sind bei überschüssiger Kapazität extrem gering. Es geht darum, schnell Größenvorteile zu erreichen.

Die Zusammenschaltung erfolgt mit Hilfe von Switchen und Routern. Bei der Erweiterung der Interconnection sind dann ggf. noch zusätzliche Portkarten oder Speicherchips hinzuzukaufen. Die Kosten hierfür sind

sprungfix. Irgendwann lässt sich ein Router nicht mehr sinnvoll erweitern und es muss ein neuer gekauft oder gemietet werden. Manche Hersteller sind bereit, das Risiko zu teilen und nutzungsorientiert abzurechnen.

Üblich ist die Zusammenschaltung in einem Telehaus. Dort kann gemeinschaftlicher Platz genutzt werden oder bei höheren Qualitätsanforderungen auch ein eigener „cage“ angemietet werden. Wenn dieser ausgelastet ist, können bei der Erweiterung erhebliche sprungfixe Kosten durch einen ggf. notwendigen Umzug entstehen, wenn umgezogen werden muss. Erhebliche Kostenelemente sind Energie inkl. Notstrom, Zugangskontrolle, Dokumentation und Feuerschutz.

Neben den unmittelbaren Kosten der Zusammenschaltung gibt es Kosten für die Nutzung der Netze. Diese Kosten werden in der Preisbildung für IP-Transit oft nur unzureichend gesehen.

Laut einer Studie von AXON (2022) verursachten die fünf größten Internet Unternehmen im Jahre 2021 über 56 % des Globalen Internet-Verkehrs, es handelt sich hier um Meta (Facebook, Instagram, WhatsApp), Alphabet (Google search, YouTube), Apple (iTunes, iCloud, AppStore), Amazon (AWS, Amazon Prime), Microsoft (MS Office, Xbox) und Netflix. Gemäß der Studie zwingen die großen Inhalte-Plattformen die Telekommunikationsfirmen mit Netzen in der Fläche zu immer neuen Investitionen und können durch ihre Verhandlungsmacht vielfach eine kostenlose Zusammenschaltung durchsetzen.

2. Formen der Zusammenschaltung

2.1 Peering

Weitgehend ohne Verträge spielt sich das sogenannte Peering ab. Hier tauschen beide Partner nur die Verkehre der jeweils eigenen Netze und der jeweils eigenen Kunden aus. In diesem Fall gibt es weder Kunde noch Provider, sondern man spricht von gleichberechtigten Peers. Beim „private peering“ werden zwei Netze unmittelbar zusammengeschaltet, sodass die Netzelemente exklusiv für die beiden Peeringpartner bereitstehen. Beim „public peering“ dagegen wird eine gemeinsame Infrastruktur auf Layer 2 Switching Plattformen genutzt.

Das public peering findet an IXPs statt. Einer der weltweit größten Internetknoten ist der DE-CIX mit weltweit 2577 verbundenen AS, einer Kapazität von 107 Tbps und 3180 Ports in Betrieb (DE-CIX, n.d.). Die Interconnection am DE-CIX gibt unmittelbar Zugang zu etwa 80 % der angeschlossenen Netzwerke, über einen sogenannten Route Server. Das ist eine

enorme Erleichterung, weil keine aufwendigen Peering-Verhandlungen geführt werden müssen, es muss nicht einmal eine Kommunikation mit dem Peeringpartner erfolgen. Unter den 20 %, die den Route Server nicht nutzen, sind aber oft wichtige Netze, am DE-CIX z.B. die Deutsche Telekom.

In der folgenden Tabelle sind die weltweit größten Austauschpunkte aufgeführt. Es sind alle, die laut Selbsteinschätzung in PeeringDB mindestens 15Tbps Internetverkehre abwickeln (PeeringDB, n.d.b).

Tabelle 2: Die 10 größten Austauschknoten weltweit

Austauschknoten	AS Anzahl
IX.br, Sao Paulo, BR	1452
DE-CIX Frankfurt, DE	967
AMS-IX Amsterdam, NL	822
NL-IX, Amsterdam, NL	397
Equinix Singapur, SG	387
ICX.br, Rio de Janeiro, BR	364
SIX Seattle, US	354
Equinix Ashburn, US	338
DATAIX, Frankfurt, DE, Stockholm, SE	302
BBIX, Broadband Internet eXchange, Tokyo, JP	252

Quelle: Eigene Auswertung der Datenbank PeeringDB.com (PeeringDB, n.d.b).

Die Zusammenschaltung ohne gegenseitige Berechnung der Kosten ist die bei weitem häufigste anzutreffende Variante, es ist das sogenannte Settlement Free Peering (SFP). Der außerordentliche Vorteil hier besteht darin, dass nur äußerst geringe Transaktionskosten anfallen, da keine Daten erfasst und weiterverarbeitet werden müssen.

Nur ganz wenige Carrier haben es geschafft, ohne Transit auszukommen und ihre gesamte Konnektivität auf SFP und den Verkauf von Transit aufzubauen. Das sind die 15 sogenannten Tier 1-Netze, die untereinander vollvermascht sind, das heißt, dass jeder mit jedem zusammengeschaltet ist.² Die Ebene der Tier 1 war zunächst allein von US-amerikanischen Carriern besetzt. Alle großen Internet Service Provider (ISP) außerhalb

2 Für eine Liste der Tier 1 Netze s. Wikipedia (Tier 1 network, 2022).

der USA waren Tier 2 Carrier, die zwar untereinander kostenlos Verkehre austauschten, aber grundsätzlich bei den US-amerikanischen Tier 1 für Transit-Verkehre bezahlen mussten.

Das Geschäftsmodell der Tier 1 Netze war zunächst sehr lukrativ. Bei der Durchleitung wurde Transit an zwei Partner verkauft, die beide bezahlen mussten, auch wenn im Extremfall die Durchleitung in einem einzigen Router geschah.

Es gibt keinerlei „Charta“ oder formelle Mitgliedschaft im Tier 1, den Status erreichte derjenige, der weltweit alle Routen ausschließlich über eigene Kunden oder kostenlose Peering bezog. Manche fassen den Tier 1 Begriff weiter als „transit-free“, da die Bedingung „kostenlos“ wegen der Vertraulichkeit der Peering-Vereinbarungen nicht transparent ist. Der einstmals führende Carrier MCI hatte in seinem Netz getrennte Router für „bezahltes Peering“ (paid peering) und „kostenloses Peering“ (SFP) eingerichtet, und damit vertrauliche Zahlungen indirekt offenbart und den Zutritt zu den Tier 1 erschwert.

Erreichen des Tier 1 Status ist nur möglich, wenn alle bisherigen Teilnehmer dem Neuling ein Peering gewähren. Je länger die Liste wird, desto schwieriger wird es, den Status zu erreichen. Die Bedingungen für das Eingehen eines Peering wird zumeist an eine Peering Policy geknüpft, exemplarisch ein Auszug der Anforderungen von Verizon (n.d.):

1. Präsenz in 25 US-Staaten, 9 Ländern in Europa, 3 Länder in Asien/Pazifik;
2. Austauschverhältnis von maximal 1,8 zu 1 – Bytes eingehend bei Verizon zu Bytes ausgehend von Verizon;
3. in den USA Mindestbandbreite von 10 Gbps, Verkehre mindestens 1,5 Gbps;
4. mindestens 1500 Transit-Kunden in den USA;
5. beide Netze verpflichten sich zur Vertraulichkeit der Bedingungen;
6. zahlreiche Anforderungen an die Qualität des Netzmanagements.

Manchmal gibt es jedoch auch beim SFP versteckte Kosten. Es ist gar nicht so ungewöhnlich, das Peering als Währung für Gegengeschäfte einzusetzen. So wird beispielsweise ein Peering als kostenlose Dreingabe beim Einkauf von Transportkapazitäten gewährt. Ein US-Carrier legte z.B. einem europäischen Netzbetreiber eine ausführliche Bilanz der gegenseitigen Geschäftsbeziehungen vor. Das Peering wurde für den Fall zugesagt, dass es gelänge, diese Bilanz auszugleichen. Die Forderung nach direkter finanzieller Kompensation ist im Internet nur schwer durchzusetzen. Viele Tier 2 Carrier fürchten, dass sie einen Präzedenzfall schaffen, wenn sie anfangen zu zahlen.

2.2 IP-Transit

Bei IP-Transit annonciert der Transitgeber den kompletten öffentlichen Adressraum des Internet. Es ist das umfassendste Angebot. Der Transitnehmer annonciert nur seinen eigenen Adressraum und den seiner Kunden, also weit weniger. Für Transit muss in aller Regel gezahlt werden.

Eine Variante ist die Bepreisung nach Anzahl der verfügbaren Routen. So gibt es etwa IP-Transit mit Limited Route Sets, zum Beispiel nur europäische Routen oder nur Routen innerhalb derselben geographischen Einheit (Stadt/Region/Land/Kontinent). Dafür gibt es dann einen geringeren Preis. Eine andere Variante erlaubt keinen Transitverkehr zu angeschalteten Peeringnetzen. Ein derartiger IP-Transit Limited Anschluss gewährt somit nur Verbindungen zu den direkt angeschlossenen Kundennetzen des anderen Netzbetreibers, nicht aber zu den Netzen der angeschlossenen Peering-AS. Bei einem solchen Limited Anschluss ist dadurch keine vollständige Erreichbarkeit des globalen Internet gegeben.

Da beim IP-Transit Geld fließt, ist er durch Verträge abgesichert. Die Käufer erhalten im Gegenzug auch Service Level Agreements. Die Abrechnung bei der Zusammenschaltung von Netzen erfolgt üblicherweise auf Basis der Kalendermonate. Die Vertragsdauer beträgt oft 1 Jahr mit automatischer Verlängerung, wenn nicht vorher fristgerecht gekündigt wird. Während der Periode großer Preissteigerungen waren auch kürzere Fristen üblich. Seit etwa 2010 sind zunehmend auch Mehrjahresverträge zu beobachten.

Eine unterschiedliche Bepreisung der Verkehre nach Internetprotokollen IPv6 oder Ipv4 ist nicht üblich, obwohl sich durch eine unterschiedliche Bepreisung u.U. ein Impuls zum Wechsel von Ipv4 auf Ipv6 erzeugen ließe. Wenn in Zukunft Ipv6 vorherrschend wird, ist es vorstellbar, dass der Zusatzaufwand bei Ipv4 extra bepreist wird.

Auf der Transportschicht könne eine Differenzierung nach den Protokollen UDP und TCP/IP erfolgen, UDP steht für User Datagram Protocol. UDP verzichtet auf hin- und her Senden sowie Überprüfen, und spart somit Bandbreite. Differenzierte Preise wurden bisher nicht beobachtet.

Im „sending party pays“ Regime zahlt derjenige, der dem Internet Daten übergibt, der sozusagen etwas „auf das Förderband“ legt (von Bornstedt et al., 2011).

Die Netzwerkelemente, die bei der Zusammenschaltung zum Einsatz kommen, erfassen SNMP (Simple Network Management Protocol) Daten, die Zähler für Bytes-In und Bytes-Out. Während bei Endkunden die Summen für Bytes-In und Bytes-Out zusammengezählt und dann tarifiert werden, ist es bei IP-Transit und bei bezahltem Peering üblich, nur den jeweils

höheren Wert zu tarifieren. Weitaus seltener genutzt ist die Abrechnung nach netflow-Daten (wie auch J-Flow, IPFIX, sFlow).

Die oben genannten Differenzen bei den Messungen auf der Lieferanten- und Kundenseite könnten zur Anfechtung der Rechnungen führen. In der Welt der Sprachtelefonie sind diese Abrechnungskonflikte (disputes) weitaus häufiger als im Internet Geschäft.

Bei IP-Transit ist es üblich, dass derjenige, der Transit bezieht, dafür bezahlt. Bei der Zusammenschaltung als Paid Peering ist es schwieriger. Manchmal erwarten beide Partner eine Bezahlung. Wenn sie sich nicht einigen können, kommt keine Zusammenschaltung zustande. Die Zahlungsbereitschaft kann sich auch an dem Wert orientieren, den der Service für die jeweils andere Seite hat. Nehmen wir einmal an, der Endkunden-provider E erzielt einen monatlichen Wertzuwachs seines Netzes durch den Videoservice V von 1,0 Mio. Euro. V erzielt durch die Interconnection mit E einen monatlichen Wertzuwachs von 1,1 Mio. Euro. Dann wäre V bereit, maximal 0,1 Mio. Euro zu zahlen. Dieses Beispiel zeigt auch, dass bei nur geringer Änderung der Parameter sich die Zahlungsrichtung umkehren kann.

3. De-facto Standards der Bepreisung

3.1 Einmalige Preise bei Zusammenschaltung

Ein solcher Preis wird erhoben zur Abdeckung der Kosten für die technischen und administrativen Tätigkeiten erhoben. Er wird auch als „non recurring charge“ (NRC) bezeichnet. Damit werden auch Einrichtungskosten einer Leitung oder das Schalten von Glasfasern in einem Telehaus abgedeckt. Der Preis richtet sich nach der Anzahl und Kapazität der gewünschten Zusammenschaltungen. Aus Wettbewerbsgründen wird gegebenenfalls auch auf einen Preis verzichtet.

Nur 3 der 942 in PeeringDB gelisteten Internetaustauschpunkte begnügen sich mit einer einmaligen Zahlung bei der Einrichtung und verzichten auf wiederkehrende Kosten, Dallas-IX, Richmond-IX und SIX Seattle (PeeringDB, n.d.a).

3.2 Laufende Bepreisung nach Anschlussbandbreite („Flat Rate“)

Flatrates sind nicht nur bei Endkunden beliebt, sondern auch bei IP-Transit Kunden. Insbesondere Kunden aus dem öffentlichen Bereich sind gezwungen, ihre Budgets einzuhalten und präferieren eine Flatrate, auch wenn sie dies letztendlich teurer kommt als eine Berechnung nach genutzten Kapazitäten.

Der monatliche Preis oder die Miete für eine Zusammenschaltung richtet sich nach der anzuschaltenden Kapazität bzw. nach der beauftragten Bandbreite. Das ist z.B. der Fall, wenn eine 100 Gbps Zusammenschaltung gebaut werden muss und der Netzbetreiber aber vielleicht ein Rate Limit von 50 Gbps eingerichtet haben möchte. Hinzu kommt eventuell der Aufwand für die Schaltung in einem Telehaus.

3.3 Laufende Bepreisung nach durchschnittlich genutzter Bandbreite („Average“)

Für diese Abrechnungsvariante werden zunächst die Zählerstände an der entsprechenden Zusammenschaltung im Router auf dem jeweiligen Interface, nach IN- und OUT-Verkehr alle 5 Minuten getrennt erfasst. Diese ermittelten Volumenzählerstände im 5-Minutentakt werden auf Bandbreite in Mbps umgerechnet. Somit erhält man für einen Monat im 5-Minutentakt 8640 diskrete Werte der genutzten Bandbreite in Mbps, getrennt nach IN und OUT. Am Ende des Abrechnungszeitraums werden die ermittelten Einzelwerte für die Bandbreite getrennt nach IN und OUT aggregiert und durch die entsprechende Anzahl der Einzelwerte, getrennt für IN und OUT geteilt, um anschließend den höheren der beiden Durchschnittswerte zur Berechnung zu verwenden.

3.4 Laufende Bepreisung nach dem 95-Perzentil

Für den Verkäufer von IP-Transit sind Belastungsspitzen kostentreibend, die Netze müssen am Maximum plus einer Sicherheitsmarge ausgelegt werden. Deshalb sind vielfach die Preise für die vertraglich zugesicherte Mindestabnahme (minimum commitment) geringer als die Preise für die erlaubten Belastungsspitzen. Üblich ist es, 10 bis 20 % der Anschlussbandbreite als Minimum commitment zu fordern.

Abbildung 1 soll verdeutlichen, wo in der Wertschöpfungskette welche Variante vorherrschend ist. Dabei ist grundsätzlich zwischen der Festnetzseite links und der Mobilfunkseite rechts zu unterscheiden. Die Stärke der Pfeile soll die Höhe der Verkehrsflüsse verdeutlichen, sie ist allerdings nicht maßstabsgerecht.

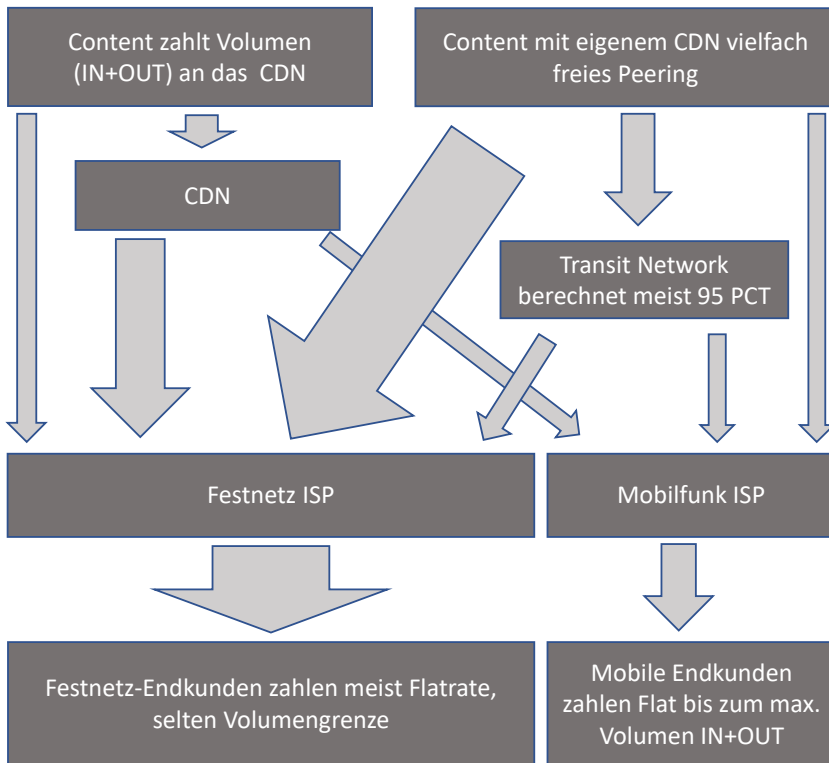
Die Endkunden zahlen für beide Verkehrsrichtungen, zwischen Unternehmen wird in der Regel nur die intensiver genutzte Verkehrsrichtung bepreist. Sofern überhaupt abgerechnet wird, hat sich die 95 Perzentil-Abrechnung als de-facto Standard bei der Zusammenschaltung im Internet etabliert. Nur wenige Transit-Provider bieten andere Prozentsätze. Beispielsweise hat der ISP Cogent auch das 90-Perzentil angeboten (Cogent, 2006).

Aus den ermittelten Werten für die Bandbreite wird pro Zählerstand der größere der beiden Werte IN oder OUT ermittelt und gespeichert. Am Ende des Abrechnungszeitraums wird die 95 Perzentil-Regel auf die gespeicherten Werte angewendet. Das bedeutet, die Werte werden nach Größe sortiert. 5 % der höchsten diskreten Werte werden gelöscht und der nächstniedrigere Wert wird als 95 Perzentil-Wert für die Preisermittlung der genutzten Bandbreite herangezogen.

Dies ist ein Kompromiss aus der Sicht des Netzbetreibers, der am liebsten den Maximalwert der Nutzung berechnen möchte. Der Verkäufer der Netzleistung muss schließlich den Maximalwert bereitstellen und sogar noch eine Sicherheitsmarge darüber hinaus. Der Käufer von IP-Transit möchte nicht den Maximalwert, sondern die durchschnittliche Nutzung bezahlen. Insbesondere möchten die Käufer nicht einen Spitzenwert bezahlen, der durch einen Hackerangriff verursacht wurde. Die 95-Perzentil-Bepreisung ist eine Form der Spitzenlastpreisbildung (peak load pricing).

Normalerweise steigt der Preis mit erhöhter Knappheit eines Gutes. Das Problem bei IP-Transit liegt nunmehr darin, dass bei allerhöchster Knappheit wegen der 95 Perzentil-Regel der Preis auf null fällt. Das Vernichten der großen 5 %-Messwerte (monatlich 36 Stunden) hat allerdings einen unerwünschten ökonomischen Effekt. Es besteht bei Kunden ein Anreiz, auf diese Spitzenbelastungen noch mehr Verkehr zu packen. Damit ist gemeint, dass man maximal 36 Stunden die Übertragungskapazität voll ausschöpfen kann, und trotzdem nur für den ansonsten niedrigen Verkehr zahlt. Wenn ein Zusammenschaltungspartner sowohl ein Champions League stream als auch ein iOS Upgrade für Apple zur Auslieferung bereithält, sollte es ökonomische Anreize geben, das Software-Update nicht während des Fußballspieles zu übertragen.

Abbildung 1: Bepreisungsvarianten auf den verschiedenen Stufen



Quelle: Eigene Darstellung.

Ein theoretisches Missbrauch-Szenario ergibt sich, wenn ein IP-Transit Kunde den Monat in 21 Zeitabschnitte bei 21 verschiedenen Providern aufteilt und damit jeden Provider nur 4,8 % eines Monats nutzt. Am Monatsende hätte er nichts zu zahlen, da er die erforderlichen 36 Stunden nicht erreicht. Dieses Extrem ist in der Realität so ausgeprägt noch nicht beobachtet worden, es mag jedoch Anreize geben, Verkehrsspitzen durch ein intelligentes Traffic Management zur Kosteneinsparung gezielt zu platzieren. Der Aufwand für diese Art von Optimierung ist bei dem geringen Preisniveau von IP-Transit jedoch zu hoch.

Darüber hinaus verhindert auch die vereinbarte Mindestnutzung einen eklatanten Missbrauch. In aller Regel wird ein sogenanntes Minimum Commitment (MC) und ein Burstanteil zur Anwendung gebracht. Ver-

traglich wird eine Mindestnutzung vereinbart, die unabhängig von der tatsächlichen berechnet wird.

In der Realität sind jedoch keine eklatanten Missbräuche der 95 Perzentil-Abrechnung bekannt geworden. Die Vorteile wären nicht langfristig, der Provider würde bei derartigem Verhalten kündigen.

Die Berechnung des Burstanteil geschieht, indem man von dem ermittelten 95 Perzentil Wert den Wert für das MC abzieht. Sollte kein Burstanteil ermittelt worden sein, so wird dem Kunden der MC in Rechnung gestellt.

Tabelle 3: Reale Werte einer Video-Distributionsplattform in Deutschland, November 2021

Durchschnittlicher Verkehr	Maximal genutzte Bandbreite	95 Perzentil
258 Gbps	772 Gbps	616 Gbps

Quelle: Basierend auf vertraulichen Daten eines repräsentativen ISP.

4. Sonderformen der Bepreisung

4.1 Entfernungabhängige Bepreisung

Im Internet spielt die Entfernung bei der Bepreisung normalerweise kaum eine Rolle, es gibt jedoch Einzelfälle. Bei der bit-Miles Verrechnung (Gilmore et al., 2016) wird die Anzahl der transportierten Bits mit der Anzahl der zurückgelegten Meilen (oder km) multipliziert. Level 3 hat im Jahr 2011 einen entsprechenden Entwurf an das FCC geliefert. 2016 hat Level 3 einen Bit Mile Peering Deal mit Google abgeschlossen (LightReading, 2016). Level 3 hat argumentiert, dass diese Art der Verrechnung fair sei. Endkunden-ISPs konnten sich dieser Meinung nicht anschließen, da die vorgelegten Bit-Miles nicht die besonderen Kosten der letzten Meile berücksichtigten.

4.2 Volumenbasierte Bepreisung

Bei Endkunden ist die volumenbasierte Abrechnung vor allem im Mobilfunkbereich zu finden. Dort werden beide Verkehrsrichtungen individuell bepreist. Bei IP-Transit zählt nur das Volumen der höheren Verkehrsrich-

tung. Diese Variante der Abrechnung ist bei IP-Transit und Peering nicht mehr so häufig. Es wird wohl noch Volumen gezählt, allerdings für ein bestimmtes Zeitintervall und dann wird das in genutzte Bandbreite umgerechnet. Damit kommen wir zu der heute gängigsten Abrechnungsvariante, die auf Bandbreite basiert.

Eine weitere Form der volumenbasierten Preisdifferenzierung sind Mengenrabatte. Bei der Zusammenschaltung sind erhebliche Mengenrabatte zu beobachten. Kleinere Netze zahlen einen mehrfach höheren Preis pro Mbps/Monat.

Tabelle 4: Mengenrabatte bei IP-Transit

Mindestabnahme	Preis per Mbps	Mindestkosten / Monat
10 Mbps	\$ 12	\$ 120
100 Mbps	\$ 5	\$ 500
1 Gbps	\$ 3,50	\$ 3500
10 Gbps	\$ 1,20	\$ 12 000
100 Gbps	\$ 0,70	\$ 70 000

Quelle: Norton (2014).

Eine seltene Preisvariante ist das „retro-active pricing“. Dort gibt es einen Rabatt ab einer bestimmten Größenordnung. Hierdurch sollen Anreize geschaffen werden, größere Volumina zu übertragen. Hierbei geschieht die Ermittlung der durchschnittlichen Bandbreite wie oben beschrieben. Allerdings wird hier dann nicht der Preis je Stufe ermittelt und diese dann addiert, sondern es gilt der Preis der höchsten erreichten Stufe für alle anderen niedrigeren Stufen ebenfalls.

Bei der inkrementellen Abrechnung gibt es den Rabatt nur jeweils für die nächsthöhere Menge. Der Verkäufer stellt sich dabei besser als beim retroaktiven Discount.

4.3 Bepreisung nach Austauschverhältnis (Ratio Billing)

Diese Art der Abrechnung findet gelegentlich bei Peering-Verbindungen Anwendung. Oft ist das Peering an ein vereinbartes Austauschverhältnis gekoppelt, das im Gleichgewicht bleiben soll. Wenn beispielsweise für das Peering ein maximales Ungleichgewicht im Austauschverhältnis von 2:1

festgelegt wurde, kann der darüberhinausgehende Verkehr zu einem bestimmten Preis verrechnet werden (vgl. z.B. Norton, 2011, S. 160).

Derartige vertraglichen Abmachungen fallen unter eine besondere Vertraulichkeit und sind bisher nicht an die Öffentlichkeit gelangt.

4.4 Ursprungs- und Zielbasierte Bepreisung

Wenn der Verkehr in Sende- oder Empfangsrichtung voneinander getrennt registriert werden kann, so ergibt sich die Möglichkeit einer individuellen Bepreisung der Verkehrsrichtungen. Somit ist dies für die Tarifierung von z.B. Videoübertragungen oder Auslieferungen von Software-Updates anwendbar, da hier der meiste Verkehr in einer Richtung übertragen wird. Der Anteil in die andere Richtung ist hierbei vernachlässigbar.

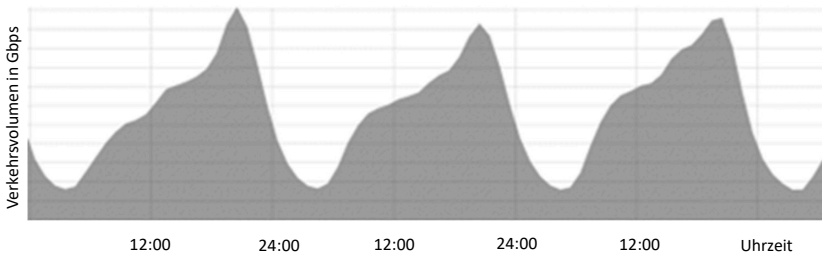
Eine solche Abrechnungsmöglichkeit erleichtert das Engpassmanagement über eine Verkehrssteuerung. Grundsätzlich kann bei dieser Abrechnungsart der Verkehr des Übergabe AS an das IP-Transit AS auf der Basis des ermittelten Ziel AS bepreist werden. Damit können bestimmte Verkehre (Verkehre der Streamingdienste, Mobilfunk, Sportübertragung usw.) billiger und andere teurer bepreist werden. Dadurch können bestimmte Routen (Quelle/Senke Verkehre) unterschiedlich monetarisiert werden. Bei mehreren Orten der Interconnection an ein IP-Transit AS kann auch der Verkehr über eine spezielle Lokation unterschiedlich bepreist und dadurch eine Verkehrssteuerung angestrebt werden.

4.5 Tageszeitabhängige Bepreisung

Eine an der tatsächlichen Auslastung orientierte Bepreisung ist im Markt noch nicht zu beobachten. Ein erster Schritt in diese Richtung könnte die Bepreisung anhand der typischen Tagesverkehrskurve sein, wie sie in Abbildung 2 gezeigt wird.

Bei der Terminierung in der Sprachtelefonie wird teilweise noch eine unterschiedliche Bepreisung zwischen Spitzenzeiten und Schwachlastzeiten als Mittel der Verkehrssteuerung eingesetzt. Vorausgesetzt wurde eine Preiselastizität der Nachfrage. Auf Kundenseite ist dies wegen zunehmender Flatrates immer weniger der Fall.

Abbildung 2: Typische Tagesverkehrskurve eines Endkundenproviders



Quelle: Basierend auf vertraulichen Daten von BENOCS (n.d.a).

Bei IP-Transit könnte auf der Basis der Tageszeit bepreist werden. Zwei Möglichkeiten sind hier zu nennen:

1. *Bepreisung nach Zahlungsbereitschaft*

Vormittags ist die geschäftliche Nutzung vorherrschend, die wenig preiselastisch ist. Demnach könnten vormittags höhere Preise verlangt werden.

2. *Bepreisung zur Reduzierung von Verkehrsspitzen*

In diesem Fall müssten die Abendstunden höher bepreist werden. So könnten elastische Verkehre wie z.B. Backups in Schwachlastzeiten verschoben werden.

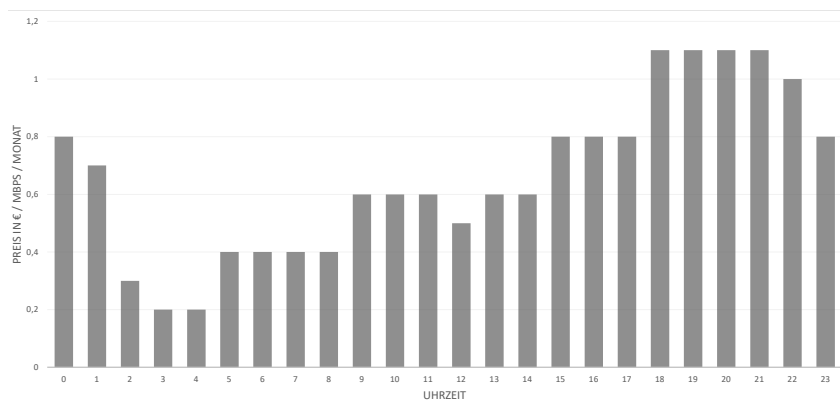
Dies soll mit einem Beispiel noch weiter erläutert werden. Ein selbstfahrendes Auto hat den ganzen Tag über umfangreiche Informationen gesammelt über seine Fahrten, diese sollen auch anderen Fahrzeugen zur Verfügung stehen. Wenn nun das Fahrzeug abends nach Hause zurückkehrt, überträgt es ausgerechnet zu Zeiten einer hohen Netzlast ein Datenvolumen an die Zentrale, das durchaus ein Terabyte groß sein kann. Netzbetreiber haben ein großes Interesse, dass dies nicht unmittelbar nach der Rückkehr geschieht, sondern erst in den späten Nachtstunden. Die Steuerung könnte sowohl an der Schnittstelle zum Endkunden als auch bei der Zusammenschaltung erfolgen. Es ist sehr viel effizienter, dies nur einmal beim Peering/Transit zu erledigen, als Millionen von Kunden das Verkehrsmanagement zur Optimierung der Netze aufzubürden.

Zusammenfassend lässt sich sagen, dass eine tageszeitbasierte Bepreisung noch nicht üblich ist bei IP-Transit. Die globale Struktur des Internet mit verschiedenen Zeitzonen kompliziert die Angelegenheit zusätzlich. Langfristig sind jedoch im Wettbewerb auslastungsabhängige Preise zu erwarten (vgl. hierzu Knieps, 2007).

5. Bepreisung nach Qualitätsparametern

Geringe Latenz ist ein wichtiges Qualitätskriterium, bei besonders hohen Anforderungen wirkt sie sich auf die Preise aus. Es gibt einen eigenen Markt für „Ultra Low Latency“ Verbindungen. Diese sind besonders kostenintensiv. Hier werden zum Beispiel Richtfunkstrecken eingesetzt, weil dort die volle Lichtgeschwindigkeit gilt und nicht die reduzierte Lichtgeschwindigkeit wie sie in Glasfasern auftritt. Glasfasern mit einem Hohlraum in der Mitte, in denen ein Vakuum herrscht, könnten ebenfalls Lichtgeschwindigkeit bieten. Diese sind jedoch noch nicht marktfähig. Ultra Low Latency Verbindungen sind um ein Mehrfaches teurer als normale Internetverbindungen.

Abbildung 3: Hypothetische Verdeutlichung Flexibler Preise in Abhängigkeit von der Tageszeit



Quelle: Eigene Darstellung.

Neben der Latenz ist die Verfügbarkeit ein wichtiges Qualitätskriterium, das auch zur Preisdifferenzierung genutzt wird. Zwei Extreme verdeutlichen das Spektrum der Kundenwünsche. Der Spieleanbieter, der unbedingt einen niedrigen Transit-Preis möchte, der es aber toleriert, wenn der Transit mal „ein paar Stunden“ nicht verfügbar ist. Auf der anderen Seite des Spektrums das Reservierungssystem der Luftverkehrsgesellschaft, die „five nines“ fordert, also 99,999 % Verfügbarkeit im Jahr, das bedeutet ein Ausfall von maximal 5,15 Minuten pro Jahr. Diese hohe Anforderung an die Verfügbarkeit erhöht den Preis nennenswert.

6. Regulierung der Zusammenschaltung

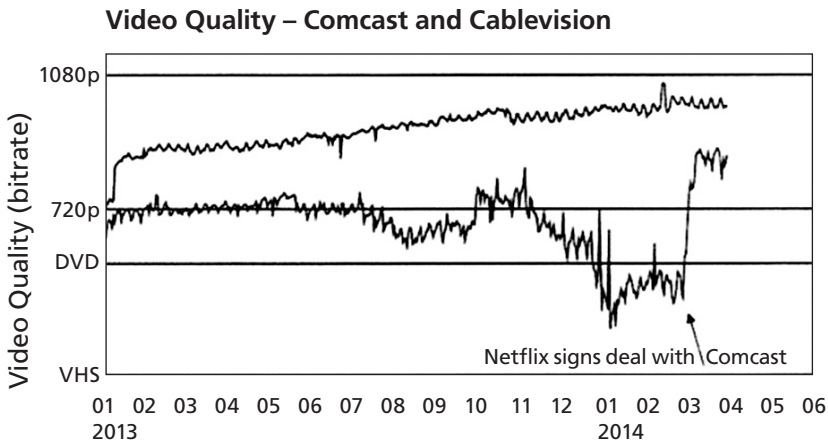
In der Sprachtelefonie ist die Regulierung der Zusammenschaltung sehr verbreitet. Für die Abrechnung der Terminierung von Telefongesprächen in andere Netze sind Preise durch Regulierer festgesetzt, im Internet haben sie sich bisher zurückgehalten, wohl aufgrund der schnellen Entwicklung wurde dies nicht für erforderlich gehalten. Die Abwesenheit von Regulierung über Jahrzehnte ist geradezu ein Erfolgsfaktor für die schnelle Innovation des Internet gewesen. Bei der hohen Rate des Preisverfalls im Internet wurde von einem funktionierenden Markt mit hoher Wettbewerbsintensität ausgegangen, der nicht der Regulierung bedarf. Peeringkonflikte werden üblicherweise über das Wettbewerbsrecht gelöst.

Bisher sind nur in der Türkei die Preise für IP-Transit reguliert. Der französische Regulierer ARCEP hat alle in Frankreich aktiven Netzbetreiber verpflichtet, quartalsweise detaillierte Statistiken einschließlich der Preise zu liefern. Der polnische Regulierer UKE hat erwogen, Peering und Transit zu regulieren.

Mögliche Anlässe für eine Regulierung könnten entstehen, wenn die Industrie Konflikte in der Zusammenschaltung nicht selbst lösen kann. Diese Streitfälle werden selten öffentlich bekannt, weil Peering- und Transitverträge meist eine Vertraulichkeitserklärung beinhalten. Öffentlich bekannt geworden ist der Konflikt zwischen dem Inhalte-Anbieter Netflix und dem Endkunden-Provider Comcast. Netflix hatte ursprünglich für die Verteilung seiner Videos DVDs per Post verschickt und dafür auch bezahlt.

Im Internet hat Netflix IP-Transit unter anderem bei Level 3 eingekauft, der Endkundenprovider Comcast musste an seinen Transit-Provider Level 3 für den Verkehr bezahlen. Comcast hatte die Kosten der Netzerweiterung in der Fläche zu tragen, ohne zusätzliche Einnahmen. Daraufhin hat Comcast Netzerweiterungen hinausgezögert, was zu drastischen Qualitätsverschlechterungen bei den Netflix-Nutzern führte. Dieser Konflikt schadete sowohl den Endkunden von Comcast als auch Comcast selbst. Letztendlich führte er wohl zu Kündigungen bei Netflix und bei Comcast. Im Gegensatz zu Europa haben Endkunden in den USA aber nur wenige Alternativen.

Abbildung 4: Qualitätsverschlechterung während des Zusammenschaltungskonfliktes zwischen Netflix und Comcast



Quelle: Seward (2014).

7. Ausblick

Trotz aller Schwächen ist die gegenwärtige 95. Perzentil-Bepreisung sinnvoll, eventuell sind höhere Werte als 95 noch besser, um die Motivation zur Abflachung von Verkehrsspitzen zu verstärken.

Das Preisniveau für die Zusammenschaltung wird vor allem in den aufstrebenden Märkten Afrika und Asien noch weiter sinken, aber wegen der großen Entfernungen auf dem Pazifik werden die Preise dort nicht ganz so stark verfallen. Es entsteht ein Wettbewerb zwischen Ländern und Regionen, die jeweils Internet-Austauschpunkte beherbergen wollen. Da ihnen selbst die nötige Erfahrung fehlt, sind sie auf renommierte Internet-Austauschpunkte wie etwa den DE-CIX angewiesen.

Die Internet-Austauschpunkte verlieren zunehmend das Geschäft mit den großen Inhalteanbietern, weil diese sich direkt mit Endkunden-ISP-s zusammenschalten. Sie sollten das kompensieren durch neue Services, die nicht nur den Austausch von Bits ermöglichen, sondern auch die Verrechnung der damit verbundenen Werte.

Die Konfiguration und das Management von Netzen werden zunehmend automatisiert. Erste Ansätze für den Einsatz von Methoden der künstlichen Intelligenz sind schon bald zu erwarten. Damit wird eine dynamische Zuteilung von Netzressourcen möglich. Im Wettbewerb sind

langfristig auslastungsabhängige Preise zu erwarten (vgl. hierzu Knieps, 2007). Potentiale zur Preisdifferenzierung können stärker genutzt werden. Die entsprechenden Abrechnungssysteme sind jetzt marktreif (vgl. z.B. die Flow-basierte Abrechnung bei BENOCS, n.d.b).

Immer mehr wird bei der Zusammenschaltung auf Sicherheitsaspekte geachtet, in dem zum Beispiel nur noch Routen akzeptiert werden, die in zertifizierten Datenbanken gelistet sind.

Die großen Cloud-Provider bauen zunehmend ihre eigenen Netze, so dass das öffentliche Internet nicht mehr ganz so stark wächst. Dadurch sinkt auch die Bedeutung der Transit-Provider und durch die sinkende Preise wird das Geschäft mit IP-Transit zunehmend uninteressanter. Die Verhandlungsmacht der großen Inhalteanbieter erzwingt Preise in der Nähe der Grenzkosten. Geoff Huston spricht seit Jahren vom „Tod des Transits“ (Huston, 2018).

Gemäß der Studie von AXON (2022) führt die Asymmetrie in der Verhandlungsmacht zwischen den großen Inhalteanbietern und den Endkundenprovidern überwiegend zu kostenlosen Zusammenschaltungen.

Durch ihre Verhandlungsmacht können sie in kommerziellen Verhandlungen mit Netzbetreibern überwiegend eine kostenlose Zusammenschaltung durchsetzen. Diese Verhandlungsmacht ergibt sich einerseits aus den starken Marktpositionen dieser Unternehmen bei den Diensten – kein Kunde würde einen Internetanschluss kaufen, wo WhatsApp Nachrichten mit Verzögerung ankommen oder YouTube Videos immer wieder stehen bleiben. Andererseits sind die Internetzugangsanbieter über die EU-Regeln zur Netzneutralität verpflichtet, (Europäische Union 2015) eine Verbindung zu allen Endpunkten des Internets herzustellen. Reziprok gibt es aber keine Verpflichtung, dass die großen Internetunternehmen ihre Dienste allen Endkunden anbieten müssen und sie sind auch nicht verpflichtet, sich diskriminierungsfrei mit allen Netzbetreibern zusammenzuschalten. Kostenlose Interconnection verringert die Anreize für die Inhalteanbieter, Netzressourcen zu sparen, wie etwa durch die Einführung effizienterer Codes die Verkehre zu reduzieren. Die Regulierer könnten versuchen, hier zu intervenieren und die Inhalteanbieter vermehrt zu einer Beteiligung an den Kosten des Netzausbaus zu drängen.

Die Grundidee des Internet war nicht geprägt von dem Konzept einer Video-Distributionsplattform, das Internet hat sich aber in diese Richtung entwickelt. Technische und ökonomische Grundannahmen müssen im Hinblick auf diese Entwicklung neu überdacht werden. Dies mag in letzter Konsequenz bewirken, dass das Internet aufgeteilt wird in eine Video-Distributionsplattform und in „traditionelles Internet“.

Referenzen

- AXON (2022). *Europe's internet ecosystem: socioeconomic benefits of a fairer balance between tech giants and telecom operators*. Axon Partners Group. <https://etno.eu/downloads/reports/europes%20internet%20ecosystem.%20socio-economic%20benefits%20of%20a%20fairer%20balance%20between%20tech%20giants%20and%20telecom%20operators%20by%20axon%20for%20etno.pdf>.
- BENOCs. (n.d.a). BENOCs Analytics. <https://www.benocs.com/analytics/> [Abgerufen am 13.02.2022].
- BENOCs (n.d.b). BENOCs 3D – Billing. <https://www.benocs.com/3d-billing/> [Abgerufen am 13.02.2022].
- Cogent. (2006, 3. April). Cogent Ups the Internet-Market Ante By Offering Burstable Bandwidth Billing at 90th Percentile Rate. <https://www.cogentco.com/de/about-cogent/press-releases/215-cogent-ups-the-internet-market-ante-by-offering-burstable-bandwidth-billing-at-90th-percentile-rate>.
- CIDR (n.d.). *CIDR Report*. <https://www.cidr-report.org/as2.0/>.
- Coll, N. (2021, 13. September). A 2021 Check-In on Global IP Transit Price Trends. *TeleGeography Blog*. <https://blog.telegeography.com/2021-global-ip-transit-price-trends>.
- DE-CIX. (n.d.). We make interconnection easy. Anywhere. <https://www.de-cix.net/de> [Abgerufen am 11.07.2022].
- Europäische Union (2015). VERORDNUNG (EU) 2015/2120 DES EUROPÄISCHEN PARLAMENTES UND DES RATES vom 25. November 2015 über Maßnahmen zum Zugang zum offenen Internet und zur Änderung der Richtlinie 2002/22/EG über den Universaldienst und Nutzerrechte bei elektronischen Kommunikationsnetzen und -diensten sowie der Verordnung (EU) Nr. 531/2012 über das Roaming in öffentlichen Mobilfunknetzen in der Union.
- Gilmore, P. W., LaPerrière, S., von Bornstaedt, F., Siegel, D., & Sejas, R. (2016, 12. April). *Bitmiles: Is there a better peering metric?*. GPF 11.0, Los Angeles, CA. https://www.globalpeeringforum.org/pastEvents/gpf11.0/presentations/Tue_3_Bitmiles.GPF11.2016-04.pdf.
- Huston, G. (2018, 18. Dezember). *The Death of Transit and the Future Internet*. Second ITU Workshop on Net2030, Hong Kong. https://www.itu.int/en/ITU-T/Workshops-and-Seminars/20181218/Documents/Geoff_Huston_Presentation.pdf.
- Knieps, G. (2007). *Netzökonomie: Grundlagen, Strategien, Wettbewerbspolitik*. Wiesbaden: Gabler.
- LightReading. (2016, 15. Januar). Level 3 Seals Bit Mile Peering Deal With Google. <https://www.lightreading.com/ethernet-ip/new-ip/level-3-seals-bit-mile-peering-deal-with-google/d/d-id/720459>.
- Norton, W. B. (2011). *The Internet Peering Playbook: Connecting to the Core of the Internet*. Palo Alto, CA: DrPeering Press.
- Norton, W. B. (2014). *What is Internet Transit?* (Excerpts from The Internet Peering Playbook: Connecting to the Core of the Internet). DrPeering International. <https://drpeering.net/core/ch2-Transit.html>.

- PeeringDB. (n.d.a). Terms: Non-recurring Fees Only [Advanced Search]. https://peeringdb.com/advanced_search?terms__in=Non-recurring+Fees+Only&reftag=ix [Abgerufen am 02.02.2022].
- PeeringDB. (n.d.b). Capacity 15Tbps [Advanced Search]. https://peeringdb.com/advanced_search?capacity__gte=15Tbps&reftag=ix [Abgerufen am 13.02.2022].
- Prince, M., & Rao, N. (2021, 23. Juli). AWS's Egregious Egress. *The Cloudflare Blog*. <http://blog.cloudflare.com/aws-egregious-egress/>.
- Seward, Z. M. (2014, 27. August). The inside story of how Netflix came to pay Comcast for internet traffic. *Quartz*. <https://qz.com/256586/the-inside-story-of-how-netflix-came-to-pay-comcast-for-internet-traffic/>.
- TeleGeography. (2020). *The State of the Network* (2020 Edition). <https://www2.telegeography.com/hubfs/assets/Ebooks/state-of-the-network-2020.pdf>.
- Tier 1 network. (2022, 5. Juli). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Tier_1_network&oldid=1094034744.
- Verizon. (n.d.). Interconnection Policy for Internet Networks. <https://enterprise.verizon.com/terms/peering/> [Abgerufen am 14.02.2022].
- von Bornstaedt, F., Röttgermann, M., Korthals, I., Johansen, F., & Lonsethagen, H. (2011, 1. Oktober). "*The Sending Party Network Pays*": *A first step towards end-to-end quality of service*. 15th International Conference on Intelligence in Next Generation Networks (ICIN 2011), Berlin. <https://doi.org/10.1109/ICIN.2011.6081074>.

Networks, Centrality, Public Good Index and the Medici

Manfred J. Holler¹ & Florian Rupp²

Abstract

This paper discusses a straightforward approach for dealing with power in networks, not taking any detour via various graph-theoretic network centrality measures: it applies the Public Good Index and analyzes the Medici marriage network in 15th century Florence with respect to power – illustrating one dimension which explains the political success of Cosimo de' Medici. It concludes with some remarks on heterogeneity due to network relations and the correspondence in spatial models in economics and politics.

Keywords:

Medici, network, centrality, power indices, Public Good Index, heterogeneity, spatial model

1. Power and the Centrality-Power Hypothesis: The Straightforward Approach

In Roger Myerson's seminal paper (Myerson, 1977) the Shapley value gives a theoretical underpinning of the bargaining solution which the author proposes for network games formalized as graphs. It also helps to identify some properties of the bargaining solution in the case that the rather specific assumptions are satisfied underlying the identity of the bargaining solution and the power measure implied by the Shapley value. Power was not the paper's concern. The focus was on presenting a bargaining solution that tells us what outcomes we should expect if the bargaining concurs with the underlying properties – and the agents are possible nodes in a network – represented by vertices in the corresponding graph. On

1 University of Hamburg and CCR-Munich, www.ccr-munich.de; manfred.holler@uni-hamburg.de.

2 Kuitaisi International University and CCR-Munich, www.ccr-munich.de; florian.rupp@kiu.edu.ge.

the other hand, network theory conceptualizes the power of the agents in a network by measuring their relationship to other members in the network. A plethora of centrality measures has been proposed (see Brandes & Hildenbrand, 2014; Todeschini & Consonni, 2009) and implemented in comprehensive software environments like R (see <https://www.r-project.org/>, especially the CINNA package), to specify this relationship. Due to their distinct nature, five of these measures can be considered most promising in view of attributing power to members of a network: *degree centrality*, *closeness centrality*, *betweenness centrality*, *eigenvector centrality*, and *diffusion centrality*.

The central assumption of *degree centrality* is that a node (i.e., a member of the graph representing the network) is important the more neighbors it has. In an undirected network, the degree of a node is the number of links (i.e., edges) this node has. Obviously, this measure does not consider the relations of the neighbors to other nodes in the network nor does it take into account the position of a node with respect to “closeness” and “betweenness” within the network – in the relation with its neighbors. *Closeness centrality* and *betweenness centrality* are proposed to cure this problem. *Closeness centrality* states that a node is important/powerful if it has a more direct influence on other vertices (see Freeman, 1978), e.g., by having better access to information at other vertices. This means that the shortest network paths to other vertices are considered. *Betweenness centrality* considers the power of a node by means of its control over, e.g., information passing between other vertices (see Brandes, 2001; Freeman, 1978). In an undirected connected network, this means that the betweenness centrality measure focuses on the extent to which a node i lies on paths between other vertices s and t . *Eigenvector centrality* assumes a node to be important if it is connected to other important vertices (see Bonacich, 1987). Mathematically, this means solving an eigenvalue problem.³ As the importance of i 's neighboring node j also depends on the importance of i , the measure becomes self-referential. The underlying calculation implies that i , when calculating its eigenvector centrality, puts a weight of 1 on its direct link with j and a weight of $\frac{1}{2}$ on the link that connects j and k if k is not connected to i otherwise – and a weight of $\frac{1}{4}$ to the link between k and r , if ... and so on. Jackson (2019, p. 10) raises the question of whether the weights 1, $\frac{1}{2}$, $\frac{1}{4}$, ..., based on a “discounting” of $\frac{1}{2}$, is adequate. Of course, the answer depends on the nature of the network, e.g., on the quality of the links. If we allow for weights that are not based on the $\frac{1}{2}$ discounting,

3 See Jackson (2010, pp. 40–43) for details.

we are likely to get a more realistic image of the network and can hope for better forecasts. This is the message and implication of *diffusion centrality*.⁴ However, how do we get the necessary information? Information is the decisive ingredient when we choose one of the centrality concepts. Obviously, diffusion centrality is a very demanding one. It “spans between two extremes of degree centrality and eigenvector centrality. If one lets the number of iterations and the probability of information moving from one node to another large enough, then diffusion centrality mimics eigenvector centrality, while with just one iteration or a tiny probability of transmission it becomes proportional to degree centrality” (Jackson, 2019, p. 10).

“What we know” is one limitation to the choice of the centrality concept, “what we want to know” is another. Whether we are interested in the efficiency of alternative network specification, from a purely theoretical point of view, or whether we want to understand how the takeover of the Medici was favored by their centrality,⁵ suggests the application of quite different concepts of centrality. The choice of the centrality concept could determine the outcome in cases centrality measures are applied to distribute network gains within a network in accordance with such measures – or allocate costs accordingly. The experiments of van Leeuwen et al. (2019) show that central players contribute significantly less than their partners in the production of a collective outcome. Here Myerson’s model of bargaining in networks (Myerson, 1977) and its correspondence to the Shapley value is relevant – at least as inspiration – bridging centrality analysis and power analysis.

An alternative question is that of the importance of decision-makers within a network when (a) the decision is on the choice of a public good and (b) the choice presupposes the support of collectivity. For instance, what environmental policy should be implemented? This kind of question illustrates the framework which underlies the following analysis. However, we will not apply centrality concepts to measure the potential impact

-
- 4 “Diffusion centrality spans between two extremes of degree centrality and eigenvector centrality. If one lets the number of iterations and the probability of information moving from one node to another large enough, then diffusion centrality mimics eigenvector centrality, while with just one iteration or a tiny probability of transmission it becomes proportional to degree centrality” (Jackson 2019, p. 10).
- 5 See Jackson (2019, pp. 33–40) and Holler and Rupp (2021, 2022) for alternative network analyses of the Medici case. Jackson’s study refers to centrality while Holler and Rupp have chosen a straightforward approach applying the Public Good Index – i.e., referring to power.

of the various agents, but their power. Quite often, centrality is a proxy for (decision) power and “important” translates into “powerful.” Jackson (2019) selected the subtitle “How Your Social Position Determines Your Power, Beliefs, and Behaviors” for his recent book “The Human Network.” For this paper, and a series of related papers,⁶ we have chosen the power dimension – and expressed power by means of a power index, i.e., we have chosen a straightforward approach dealing with power in networks, not taking any detour via centrality.

More specifically we applied the Public Good Index to capture the possible impact of a member of the considered society, i.e., an agent (or player) on the selection of a specific public good (from the pool of all possible public goods). Referring to the possible impact, we consider power as a potential. This concurs with Max Weber’s definition of power: “In general, we understand by ‘power’ the chance of a man or of a number of men to realize their own will in a communal action even against the resistance of others who are participating in the action” (Weber, [1924] 1948, p. 180).⁷

The following section presents an introduction of the Public Good Index and the, closely related, Public Value. Section 3 describes the Medici marriage network of medieval Florence which will serve as input to a power analysis of this very network, presented in section 4, in order to illustrate the working of the power measures. Section 5 concludes the paper with some remarks on heterogeneity.

2. Public Good Index (PGI) and Public Value (PV)⁸

The Public Good Index (PGI) and the Public Value (PV) refer to minimum winning coalitions (MWCs), i.e., coalitions which are just large enough to produce a particular outcome. For instance, if C is a winning coalition, then it is minimal if $C \setminus \{i\}$ is a losing coalition, i.e., any member i leaving coalition C turns the winning coalition C into a losing coalition $C \setminus \{i\}$. (C and $C \setminus \{i\}$ are set of players of the corresponding cooperative game.)

6 Holler & Rupp (2019, 2020, 2021, 2022).

7 Note that in the German original Max Weber used “Chance” which expresses a possibility or a potential, and not a probability. However, Talcott Parsons translated “Chance” by “probability” (Weber 1947, p. 152) – he should have known better.

8 This material derives from Holler and Rupp (2020). See also Holler and Rupp (2021, 2022).

If C is a winning coalition and $C \setminus \{i\}$ is a losing coalition then i is essential for the winning of C , i.e., player i is critical; i is a swing player; i has power.⁹

The PGI has been introduced in Holler (1982) and axiomatized in Holler & Packel (1983).¹⁰ It is a vector representing the numbers of MWCs which have nodes i for an element. For a particular node i this number is c_i . The numbers c_i get standardized such that the “shares” of all nodes add up to one. Thus, $h_i(\Gamma, d)$, the PGI of i given a particular network structure Γ and a decision rule d , is

$$h_i(\Gamma, d) = \frac{c_i}{c} = PGI_i(\Gamma) \text{ with network } \Gamma \text{ and } \sum_{i=1}^n c_i = c \quad (1)$$

Here, c_i is a function of Γ and d . In the case of a collective decision problem d is the decision rule, e.g., a majority quorum.

The recipe for calculating the PGI, therefore, is: (a) We list the minimum winning coalitions, (b) count for each agent i how many minimum winning coalitions have i as a member, (c) sum up the resulting c_i counts, and (d) divide each c_i count by the sum of this count. Step (d) gives us ratios which, of course, add up to one. As Holler and Li (1995) demonstrated, by looking at the ratios only, relevant information can be lost. Therefore, we will also present the non-standardized numbers, i.e., the Public Value (PV): the PV of i is identical with the number of MWCs which have i as a member.

In order to illustrate the application of the PGI to networks, we will discuss a toy example demonstrating that you do not need to have resources to have some power when you are connected. Fragnelli (2012) offers a simple example that demonstrates the power of a link.¹¹ The decision body is given by the weighted voting game $v^* = (51; 35, 30, 25, 10)$, i.e.,

-
- 9 The swing players define the power for the indices of Shapley-Shubik, Penrose-Banzhaf-Coleman, Johnston and Deegan-Packel. These measures vary in what coalitions are to be considered and how the coalition surplus is distributed among their members.
 - 10 There are alternative axiomatizations in, e.g., Alonso-Mejide et al. (2008) and Safokem et al. (2021). For a recent discussion of the PGI, see Holler (2019), Kurz (2020), and Soh Voutsas (2021). For extensions, see Courtin & Tchantcho (2019) and Kurz (2021).
 - 11 However, Fragnelli (2012) applies the Banzhaf index to measure power which takes into account all winning coalitions. It is not obvious that this index expresses

the decision rule and vote distribution are $d = 51$ and $w^* = (35, 30, 25, 10)$, respectively. Note that player 4, having a weight $w_4^* = 10$, is a *dummy* in v^* ; player 4 is not *critical* for the winning of any coalition. The set of MWCs is $\{\{1,2\},\{1,3\},\{2,3\}\}$. The corresponding power distribution, measured by the PGI, therefore is

$$b(v^*) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{0}{3}\right).$$

The voting weights can be interpreted as proxies for the resources an agent can invest into the production of a public good. Solon divided Athens' citizenship into four property-classes. Belonging to a given class does not constitute a "right" describing what you can get from your community, but an expectation which the community had of a particular member, i.e., what he can contribute. In principle, this still applies to paying taxes.

The result $b(v^*)$ implies that all minimum winning coalitions are equally likely and there are no specific relationships between any pair of players in game v^* – more specifically, that there is no underlying network. To see the possible effect of such a network structure, let us add a linear network structure, $\Gamma^\circ: (1) - (2) - (3) \dots (4)$, to the weighted voting game $v^* = (51; 35, 30, 25, 10)$, with player 4 being unconnected. The set of MWCs is $\{\{1,2\},\{2,3\}\}$. The corresponding power distribution therefore is

$$b(\Gamma^\circ) = \left(\frac{1}{4}, \frac{2}{4}, \frac{1}{4}, \frac{0}{4}\right).$$

Player 2's dominating power is the result of centrality as the voting weights of players 1, 2, and 3 represent equal voting power if we abstract from the network. The impact of the network positions on the power distribution is even more obvious when we change the network structure to $\Gamma': (1) - (4) - (2) - (3)$. The set of MWCs and the values of the PGI are $\{\{1,4,2\},\{2,3\}\}$ and

$$b(\Gamma') = \left(\frac{1}{5}, \frac{2}{5}, \frac{1}{5}, \frac{1}{5}\right).$$

The example demonstrates that *we do not need resources if we are well connected*. This looks like a good message to the have-nots. However, many people are well connected only because of their resources.

the impact of a decision maker in the case of a public good. (see Holler, 1982, 2019).

There are more than ten (or twenty) power indices, but there are only two measures that are serious about measuring the impact of a player in producing a *public good*, i.e., the PGI and the PV. Because of public good assumption, there is “no splitting up of a cake” and no bargaining over shares as in Myerson (1977) and the contributions that build on it. Of course, there are networks producing (private) goods that invite sharing, but the power of a family was not shared in a Medici world – it was a collective good for the family. Padgett and Ansell (1993) studied the development of the networks of the important families of Florence for the period 1395–1434 with respect to centralization with a focus on the Medici. They concluded that the dimensions of patronage and friendship do not really contribute to explaining political partisanship; they cannot be considered “legitimate independent predictors” (Padgett & Ansell, 1993, fn.28). Both the marriage and the business network defined the power relations in the City of Florence in the first half of the 15th century. In this note, we will restrict ourselves to the power index analysis of the marriage network of the Medici family.¹² As a consequence, what follows is a partial analysis.

3. The Medici Marriage Network

Figure 1¹³ is a stylized representation of the Medici marriage network; it derives from the seminal work of Padgett and Ansell (1993). The nodes represent the elite families relevant for Florence politics in the first half of the 15th century (numbers are assigned in alphabetical order). The marriage relations between the families are mapped by the edges, i.e., the links, which connect the nodes. Nodes and edges define the network which illustrates the marriage relations.

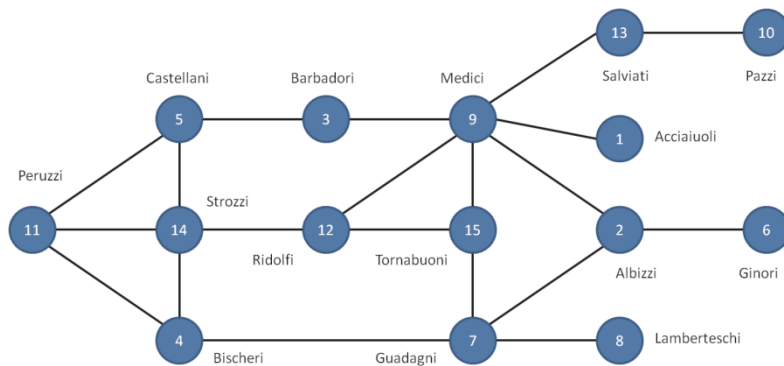
The results in Padgett and Ansell (1993, p. 1285) indicate that the oligarchs, i.e., the contenders to the Medici party, were “densely interconnected, especially through marriage.” The Medici, on the other hand, intermarried, “nearly exclusively,” with patricians outside their own neighborhood. Only “14 % of Medici marriages were within their home quarter, compared with 36 % for the superelite.” Padgett and Ansell (1993, p. 1298) observed that exogamy was “forced on the Medici by the resurgent olig-

12 Holler & Rupp (2021) consider both the marriage and the business network of the Medici.

13 Figure 1 and Table 1 derive from Holler & Rupp (2021).

archs' successful blockage of Medici marriage with their San Giovanni patrician neighbors." But it turned out that this blockage policy was rather favorable for the Medici; the network analysis shows that it is quite likely they benefitted from being discriminated against. An extended power index analysis could confirm this hypothesis. One of the roots of the blocking policy was the Ciompi revolt – the rebellion of the wool workers – of 1378 which was supported by Salvestro de' Medici. Although the latter was a sideline member of the Medici family, the Medici were shunned by the aristocratic families, especially by the families of their neighborhood.

Figure 1: Florentine elite marriages 1395–1434¹⁴



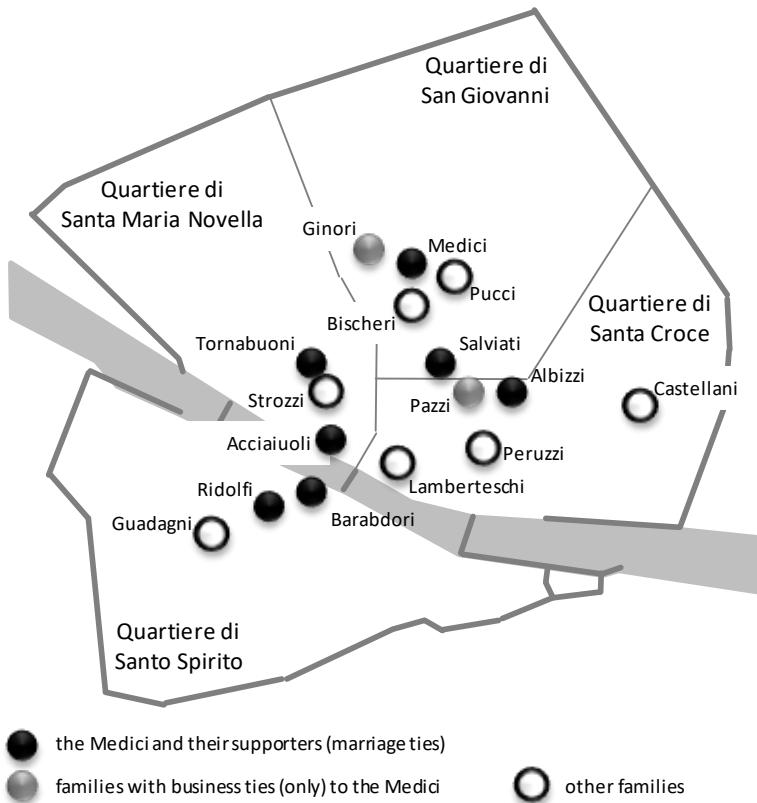
Source: Author's representation based on data by Schoch (2017)

The Ciompi revolt of 1378 was a rebellion among unrepresented laborers who did not belong to any guilds and who were, therefore, unable to participate in the Florentine government. Through their major guilds, the merchants quickly realigned with artisans in minor guilds to reestablish control, but in 1382 this alliance was overthrown by an embittered patrician elite which ruled Florence until the Medici took over in 1434 (Padgett & Ansell 1993). Much of the Florentine network structure, outlined in Figure 1, is the result of the Ciompi revolt and how the various patrician families sympathized with it. Throughout the decades following the Ciompi revolt, the leaders of the Medici clan tried hard to distance themselves from their more fractious relatives and to reestablish their family within the conservative oligarchic regime – without success. The result is a rather

14 The Pucci family is not considered in the network scheme because this family had no marriage links to any other family in this network.

diverse marriage network in which the Medici take up a central position. Padgett and Ansell (1993, p. 1287) concluded that the contending “oligarchs crafted the networks of their own destruction.”

Figure 2: *Geographic location of the Medici's connections through marriage and business ties within the four quarters of the medieval city of Florence¹⁵*



Source: Own presentation based on Fink (2016).

Where did these families live? Where were their homes in the City of Florence? The Peruzzi lived in Santa Croce, and so did the Ricasoli, and Castellani. The Strozzi and Gianfigliuzzi were the leading families

15 Based on Simon Fink's (2016) geographic location study of the headquarters of the super-elite families.

in Santa Maria Novella. The Peruzzi, Albizzi, Strozzi, and Gianfigliuzzi were classified as superelite families. The Medici – who, by the rating of the contending oligarchs, were considered members of the party of the people – lived in the western part of San Giovanni, while the Albizzi (and Guadagni) lived in the eastern part, see Figure 2.

Back to Figure 1. It illustrates that the Medici had a direct marriage link to the Albizzi, which is perhaps surprising because Rinaldo degli Albizzi was the main challenger of Cosimo – in fact, in 1433, he wanted to see Cosimo killed instead of being exiled. Figure 1 does not show the intensity of the various links: it does not express how many marriages there were, covered by a link, and how recently the link was created. Perhaps the intensity did not matter and there was no discounting. Table 12 in Padgett and Ansell (1993, p. 1299) structures the marriages into four time periods. But numbers are low, and, in many periods, there were no marriage links between two particular elite families at all. Given this, the high level of abstraction of Figure 1 seems quite acceptable. Interestingly, it shows elements of a star network system with the Medici position as a hub, but also of extended network chains with Medici links. We should expect that this has an impact on forming alliances – perhaps parties.

4. Medici Power in the Marriage Network

To study this issue and to get a better understanding of the marriage network, on the one hand, and the implications of the power index analysis, on the other, we evaluated the power distribution within the network – we asked: how large is the potential of each family to form alliances with other members of the marriage network, given the relations between the families as expressed by Figure 1? We calculated the Public Good Index (PGI) and Public Value (PV) to get an approximative answer to this question. Table 1 shows the values of the PGI, the PV, and the ranking of the marriage relations of 15 elite families in 15th century Florence in accordance with these values.

The underlying decision model, which is at the heart of our power analysis, is a voting game $v^o = (d, w)$ which assigns a voting weight $w_i = 1$ to each family i and a decision rule $d = 8$, i.e., simple majority: $v^o = (8; 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$. As suggested by the recipe above, we wrote down all minimum winning coalitions – *taking care of the network structure* defined by Figure 1 – and counted how often family i was included. This gave us the value c_i which is identical with the PV of i . After standardization, in accordance with (1) we obtained the corresponding PGI value. The

calculations were done with the help of an appropriate computer program written in the programming environment R (see <https://www.r-project.org/>).

Table 1: Public Good Index (PGI), Public Value (PV), and ranking of the marriage relations of 15 elite families in the 15th century (sorted according to the PGI-ranking)

	PGI	PV	Rank
9 Medici	0,11787879	778	1
7 Guadagni	0,09136364	603	2
2 Albizzi	0,08787879	580	3
4 Bischeri	0,07348485	485	4
14 Strozzi	0,07303030	482	5
12 Ridolfi	0,07272727	480	6
3 Barbadori	0,07212121	476	7
13 Salviati	0,06969697	460	8
15 Tornabuoni	0,06560606	433	9
5 Castellani	0,06469697	427	10
1 Acciaiuolli	0,05196970	343	11
11 Peruzzi	0,05606061	370	12
8 Lamberteschi	0,03893939	257	13
6 Ginori	0,03742424	247	14
10 Pazzi	0,02712121	179	15

It could be argued that the assumption of equal weights $w_i = 1$ to each family is too simplistic, perhaps trivial, and surely does not concur with reality. On the other, because of this assumption, we can argue that the differences in the power values are the result of the positions and relations of the various agents in the network, and not due to differences in the resources. This concurs with the implications of the toy models in section 2.

The numerical results, listed in Table 1, demonstrate the outstanding position of the Medici family in this network – followed by the Guadagni and the Albizzi, two families who lived in the same area of Florence and intermarried. From this point of view, it does not come as a surprise that the Albizzi were the rivals of the Medici. In fact, during his second

term, the elected Gonfaloniere Bernardo Guadagni (1367–1434) arrested, imprisoned, and finally sent Cosimo de' Medici and a larger number of his family members and followers into exile – in order to satisfy an initiative by Rinaldo degli Albizzi to destroy the Medici family.

5. *Network Heterogeneity*

The toy models in section 2 and the implicit voting model underlying the power analysis of the marriage model demonstrated that, in general, networks create heterogeneity if they lack symmetry – even the resources of the agents are identical. Of course, this does not come as a surprise. In fact, spatial models – which, as we see above, are platforms for representing networks – have been designed to model heterogeneity of commodities and services but also in the political space. In his pioneer paper, Harold Hotelling assumed two vendors that compete in a linear space “which may be Main Street in a town or a transcontinental railroad.” Transportation costs are not negligible depending on the distance of the buyer to the location of the supplier. Even if we assume identical production costs and homogeneous products, goods will be heterogeneous due to the locations which determine two seller-buyer networks in this case (if sellers do not move), and sellers can realize non-identical prices despite the homogeneity of the products.

Hotelling also applied his model to the political market: “The competition for votes between the Republican and Democratic parties does not lead to a clear drawing of issues, an adoption of two strongly contrasted positions between which the voter may choose. Instead, each party strives to make its platform as much like the other's as possible. Any radical departure would lose many votes, even though it might lead to stronger commendation of the party by some who would vote for it anyhow. Each candidate ‘pussyfoots,’ replies ambiguously to questions, refuses to take a definite stand in any controversy for fear of losing votes. Real differences, if they ever exist, fade gradually with time though the issues may be as important as ever” (Hotelling 1929, pp. 54f.). This became the Downsian model of Public Choice theory although it seems that Antony Downs (1957) was neither aware of Hotelling's article nor of Duncan Black's median voter equilibrium (Black 1948) foreshadowed by Hotelling. The latter says: given two candidates (Republicans and Democrats) and a left-right continuum of the voters, both candidates will choose the median voter position, i.e., offer identical platforms and share votes equally.

Needless to say that it is straightforward to conclude that the candidates had equal power if they were free to move in this one-dimensional space, and not slowed down by some ideological presumption. But even then the power question gets more complicated, also because an equilibrium in pure strategies may not exist.

Soh Voutsas (2021) offers a spatial model. The fundamental assumption of this model is the independence of the voters, i.e., “allowing voters to freely choose their own ideal points in the ideological space, consequently and in contrast to some existing spatial models like Shapley (1977), gives to voters the right to have same ideal points (which oftentimes occurs in real political situations)” (Soh Voutsas 2021, p. 25). For Soh Voutsas’s spatial model, he could (so far) offer only two indices that are consistent with the independence assumption: the Banzhaf index and the PGI while this is, e.g., not the case with the Shapley-Shubik index. On the one hand, this is a good message when we use the PGI, on the other, the space defined by the Florentine networks does not offer independence to the agents. On the contrary, it is defined by various rigidities – as a result of business relations, marriages, and the historical past. There is some research waiting for us linking the properties of the networks to the properties of the various power measures.

References

- Alonso-Meijide, J. M., Casas-Méndez, B., Holler, M. J., & Lorenzo-Freire, S. (2008). Computing power indices: Multilinear extensions and new characterizations. *European Journal of Operational Research*, 188(2), 540–554. <https://doi.org/10.1016/j.ejor.2007.04.019>
- Black, D. (1948). On the Rationale of Group Decision-making. *Journal of Political Economy*, 56(1), 23–34.
- Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5), 1170–1182.
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology*, 25(2), 163–177. <https://doi.org/10.1080/0022250X.2001.990249>
- Brandes, U., & Hildenbrand, J. (2014). Smallest graphs with distinct singleton centers. *Network Science*, 2(3), 416–418. <https://doi.org/10.1017/nws.2014.25>
- Courtin, S., & Tchantcho, B. (2019). Public Good Indices for Games with Several Levels of Approval. *The B.E. Journal of Theoretical Economics*, 20(1), 20190068. <https://doi.org/10.1515/bejte-2019-0068>
- Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper.

- Fink, S. (2016, May 11). The Medici Marriage Network in Florence. *Simon Fink*. <https://simonfink.wordpress.com/2016/05/11/the-medici-marriage-network-in-florence/>
- Fagnelli, V. (2012). A Note on Communication Structures. *Homo Oeconomicus*, 29(2), 163–171.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- Holler, M. J. (1982). Forming Coalitions and Measuring Voting Power. *Political Studies*, 30(2), 262–271. <https://doi.org/10.1111/j.1467-9248.1982.tb00537.x>
- Holler, M. J. (2019). The Story of the Poor Public Good Index. In N. T. Nguyen, R. Kowalczyk, J. Mercik, & A. Motylska-Kuźma (Eds.), *Transactions on Computational Collective Intelligence XXXIV* (pp. 171–179). Heidelberg & Berlin: Springer.
- Holler, M. J., & Li, X. (1995). From public good index to public value: An axiomatic approach and generalization. *Control and Cybernetics*, 24(3), 257–270.
- Holler, M. J., & Packel, E. W. (1983). Power, Luck and the Right Index. *Zeitschrift Für Nationalökonomie / Journal of Economics*, 43(1), 21–29.
- Holler, M. J., & Rupp, F. (2019). Power in Networks: A PGI Analysis of Krackhardt's Kite Network. In N. T. Nguyen, R. Kowalczyk, J. Mercik, & A. Motylska-Kuźma (Eds.), *Transactions on Computational Collective Intelligence XXXIV* (pp. 21–34). Heidelberg & Berlin: Springer. https://doi.org/10.1007/978-3-662-60555-4_2
- Holler, M. J., & Rupp, F. (2020). Power in Networks and the Urban Space. In E. Macri, V. Morea, & M. Trimarchi (Eds.), *Cultural Commons and Urban Dynamics A Multidisciplinary Perspective* (pp. 37–52). Basel: Springer International Publishing.
- Holler, M. J., & Rupp, F. (2021). Power in Networks: The Medici. *Homo Oeconomicus*, 38(1–4), 59–75. <https://doi.org/10.1007/s41412-021-00108-1>
- Holler, M. J., & Rupp, F. (2022). Power in Networks, the Urban Space, and the Medici. *Munich Social Science Review*, 5, 21–32.
- Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 39, 41–57. <https://doi.org/10.2307/2224214>
- Jackson, M. O. (2010). *Social and Economic Networks*. Princeton, NJ: Princeton University Press.
- Jackson, M. O. (2019). *The human network: How your social position determines your power, beliefs, and behaviors*. New York, NY: Pantheon Books.
- Kurz, S. (2020). Which Criteria Qualify Power Indices for Applications? – A Comment on “The Story of the Poor Public Good Index.” *Munich Social Science Review*, 3, 185–198.
- Kurz, S. (2021). *The Public Good index for games with several levels of approval in the input and output* (arXiv:2107.10458 [cs]). <http://arxiv.org/abs/2107.10458>
- Myerson, R. B. (1977). Graphs and Cooperation in Games. *Mathematics of Operations Research*, 2(3), 225–229. <https://doi.org/10.1287/moor.2.3.225>

- Padgett, J. F., & Ansell, C. K. (1993). Robust Action and the Rise of the Medici, 1400–1434. *American Journal of Sociology*, 98(6), 1259–1319.
- Safokem, A., Moyouwou, I., & Mekuko, A. Y. (2021). Axiomatizing the Public Good Index via merging and new arrival properties. *Discrete Applied Mathematics*, 305, 86–102. <https://doi.org/10.1016/j.dam.2021.08.024>
- Schoch, D. (2017). *netrankr: An R package to analyze partial rankings in networks*. <https://CRAN.R-project.org/package=netrankr>
- Shapley, L. S. (1977). *A Comparison of Power Indices and a Nonsymmetric Generalization* (P-5872; The Rand Paper Series). RAND Corporation. <https://www.rand.org/pubs/papers/P5872.html>
- Soh Voutsas, A. C. (2021). *The Public Good spatial power index in political games* (THEMA Working Paper n°2021–01). CY Cergy Paris Université.
- Todeschini, R., & Consonni, V. (2009). *Molecular descriptors for chemoinformatics: Vol. Volume I&II* (2nd ed., revised and enlarged). Weinheim: Wiley-VCH.
- van Leeuwen, B., Ramalingam, A., Rojo Arjona, D., & Schram, A. (2019). Centrality and cooperation in networks. *Experimental Economics*, 22(1), 178–196. <https://doi.org/10.1007/s10683-018-9592-1>
- Weber, M. (1947). *The Theory of Social and Economic Organization* (T. Parsons, Ed.). New York, NY: Free Press.
- Weber, M. ([1924] 1948). Class, Status and Party. In H. H. Gerth & C. W. Mills (Eds.), *Essays from Max Weber*. London: Routledge & Kegan Paul. <https://doi.org/10.4324/9780203759240>

Lessons learned – Anstöße der Regulierung zum Abschreibungsproblem

Hans-Ulrich Küpper¹

Abstract

Günter Knieps und ich haben Ende der 90er Jahre im Hinblick auf die Ermittlung kostenorientierter Netzentgelte das Abschreibungsproblem neu analysiert. Der Prozeß der Regulierung von Netzindustrien hat weltweit, für mich vor allem in der Diskussion mit Stefan Reichelstein, Stanford, zu neuen Einsichten geführt. Dieser Weg wird hier in seinen wichtigsten Schritten nachgezeichnet. Damit verbunden ist der Wunsch, die Ergebnisse sollten Eingang nicht nur in die Lehrbuchliteratur finden, sondern auch von der Praxis im internen wie im externen Rechnungswesen umgesetzt werden.

Keywords:

Abschreibungen, Annuitätenabschreibung, Anschaffungswerte, Netzindustrie, Wiederbeschaffungswerte

1. (De-)Regulierung der Monopole in Netzindustrien als Anlass der Untersuchung

Aus heutiger Sicht deutet vieles darauf hin, dass sich das Jahr 1990 als Einschnitt und Wendepunkt herausstellen wird. Das gilt vor allem für Politik und Geschichte, durch den Fall der Mauer in Berlin, den Zusammenbruch der kommunistischen Ordnung in Osteuropa und die hierdurch beflügelte Globalisierung. Aber auch für die Wirtschaftswissenschaften kann man ab 1990 eindruckliche Veränderungen erkennen. An den Hochschulen gewannen wirtschaftliche Überlegungen und Instrumente mehr Bedeutung

1 Prof. Dr. Dr. h.c. Hans-Ulrich Küpper, Institut für Rechnungswesen, Strategie und Innovation der Universität Wien; Direktor i.R. des Instituts für Produktionswirtschaft und Controlling, Ludwig-Maximilians-Universität München; kuepper@lmu.de.

und wurden, zuerst von vielen kaum wahrgenommen, manche nach 1968 auf Demokratisierung ausgerichtete Strukturen im Hinblick auf eine höhere Leistungsfähigkeit und Effizienz verändert (vgl. Küpper, 2009b). Für die Bewertung der Forschung erhielten internationale Zusammenarbeit und Veröffentlichungen ein deutlich höheres Gewicht. In der Wirtschaftspraxis gewannen institutionelle Investoren an Bedeutung und mit ihnen setzte sich eine stärkere Beachtung von Shareholder Value-Konzepten sowie theoretischen, insbesondere kapitaltheoretischen Konzepten durch.

Gleichzeitig strebte die Europäische Union durch die Öffnung von Märkten einen Abbau von Monopolen (vgl. hierzu Knieps, 2007) an. Bevor Marktbedingungen erreichbar waren, mussten dazu Unternehmungen in monopolistischen Engpassbereichen wie die Deutsche Telekom AG einer Regulierung unterworfen werden. Dieser Prozess erforderte von ihnen große Anstrengungen, zu dem beispielsweise der Aufbau eines entsprechenden Rechnungswesens gehörte, um die zur Genehmigung regulierter Preise erforderlichen Daten zu erfassen und der Regulierungshörde bereit zu stellen.

In diesem Prozess sind Günter Knieps und ich uns begegnet. Wir beide sollten diesen bei der Deutschen Telekom AG begleiten. Dies hat zu einer sehr schönen Zusammenarbeit geführt, die ihren Niederschlag in einem daraus folgenden Aufsatz in der Zeitschrift für betriebswirtschaftliche Forschung gefunden hat (Knieps et al., 2001).

Für mich war das überraschendste Phänomen, welche Bedeutung Konzepte, Erkenntnisse und Instrumente der Kostenrechnung wieder gewonnen haben (Küpper, 2002). Sie gehört zu den ältesten Gebieten der deutschen Betriebswirtschaftslehre.² Zwar war sie immer ein unabdingbarer Teil jedes BWL-Studiums von der kaufmännischen Lehre bis zum Universitätsstudium. Jedoch wirkte sie mit ihren Verfahren z.B. der Zuschlagskalkulation eher ‚altertümlich‘. Durch die Regulierung benötigte man sie zur Begründung von Preisen in regulierten Netzindustrien, und es wurden Begriffe wie Einzel- und Gemeinkosten sowie Prinzipien und Verfahren der Kostenverteilung wichtig, deren Folgen hohe (Millionen-)Beträge ausmachen konnten. Dies war vor allem darauf zurückzuführen, dass zumindest in der ersten Phase der (De-)Regulierung kostenorientierte Preise im Vordergrund standen.

Tiefer gehend ist für mich die Erkenntnis, dass die Auseinandersetzung mit den durch die (De-)Regulierung zu lösenden Fragen den Anstoß dafür gab, zentrale Fragen der Kostenrechnung neu zu durchdenken und auf

2 Vgl. z.B. Schmalenbach, E. (1925), Mellerowicz, K. (1933), Kosiol, E. (1964).

theoretischer Basis zu aus meiner Sicht neuen und wesentlich besseren Konzepten für ihre Lösung zu gelangen. Deshalb finde ich es interessant, und in einer Festschrift für Günter Knieps passend, die maßgeblichen Schritte auf dem Weg zu ihnen zusammenzustellen. Für mich sind sie ein Zeugnis dafür, wie wertvoll unsere Zusammenarbeit vor 25 Jahren war.

Nach einem Verweis auf die traditionelle Behandlung der Abschreibung in Abschnitt 2 gehe ich in Abschnitt 3 darauf ein, wie die Regulierung den Anstoß zu seiner erneuten Analyse gab. Abschnitt 4 untersucht die wichtigsten Alternativen und neuen Erkenntnisse zu seiner Lösung, während Abschnitt 5 deren Bedeutung durch die wirklichkeitsnähere Betrachtung eines ausgeglichenen Anlagenbestands neu einordnet. Der Beitrag endet in Abschnitt 6 mit der Aufforderung zur Umsetzung der Erkenntnisse in BWL und Praxis.

2. Abschreibungen als ‚altes‘ Kernproblem des Rechnungswesens

Abschreibungen spielen im externen wie im internen Rechnungswesen eine zentrale Rolle. Über plan- und außerplanmäßige Abschreibungen im Anlage- sowie im Umlaufvermögen wird der in der Gewinn- und Verlustrechnung ermittelte und im Jahresabschluss ausgewiesene Erfolg einer Unternehmung beeinflusst. Während beispielsweise beim Einsatz von Material und Personal die dafür geleisteten Auszahlungen oft mit den als Aufwendungen in der externen bzw. als Kosten in der internen Rechnung angesetzten Beträgen übereinstimmen, weichen die Auszahlungen z.B. für Gebäude und Anlagen deutlich von den einer Periode zuzurechnenden Abschreibungen ab. An diesen lässt sich der Unterschied zwischen einer Zahlungs- (Kapital-, Geldfluss- oder Cash Flow-) und einer Periodenerfolgsrechnung am klarsten aufzeigen. Der (Ein-)Zahlungsüberschuss einer Periode sagt nichts über den in ihr erzielten ‚Erfolg‘ aus, weil ein hoher Überschuss auf zu geringe Investitionen und ein Überschuss von Null auf die Vornahme zweckmäßiger Investitionen hinweisen kann. Abschreibungen sind daher ein zentrales Bindeglied zwischen der auf die Liquidität ausgerichteten Finanzrechnung und den auf den Periodenerfolg ausgerichteten Rechnungen der Bilanz- und der Kosten- sowie Erlösrechnung.

Traditionell wurden Abschreibungen z.B. als „der durch einen Geldbetrag erfasste Werteverzehr des Anlagevermögens“ (Schweitzer et al., 2016, S. 116) definiert. Man fragte nach den Ursachen der Wertminderung, die insbesondere im Verschleiß, einem Fristablauf, einer Überholung, und Werteinbußen bzw. Wertminderung gesucht wurden (vgl. u.a. Schweitzer et al., 2016, S. 116ff.). Die Bemessung periodischer bilanzieller im externen

Rechnungswesen oder kalkulatorischer Abschreibungen im internen Rechnungswesen orientiert sich an ihrem Ausgangswert, der Abschreibungssumme, dem Abschreibungszeitraum und dem Abschreibungsverfahren. Dabei liegt ein wichtiger Unterschied zwischen externem und internem Rechnungswesen darin, dass man in ersterem insb. nach § 253 Abs. 1 HGB von den Anschaffungs- und Herstellungskosten, im externen dagegen traditionell von den Wiederbeschaffungskosten ausgeht. Charakteristisch für beide Rechnungen ist der Blick zurück – maßgeblich sind die bei der Anschaffung in der Vergangenheit bzw. die gegenwärtig geltenden Werte.

Da Abschreibungen eine Schnittstelle zwischen lang- und kurzfristiger Sicht bilden, werden ihre Funktion und Bestimmung seit langem intensiv diskutiert. Als planmäßige Abschreibungen bilden sie den Einsatz längerfristig genützter Güter ab, deren Anschaffung auf Entscheidungen beruht, die üblicherweise mit Methoden der Investitionsrechnung fundiert werden. Sie gehen aber in die kurzfristig ausgerichteten bilanziellen bzw. kalkulatorischen Rechnungen ein. Deshalb liegen ihnen immer Hypothesen zugrunde, über welche die Verknüpfung zwischen beiden Planungshorizonten hergestellt wird (Küpper, 1984, 1985).

3. Der Anstoß zum Aufrollen des Abschreibungsproblems durch die Regulierung

Die nach 1990 einsetzende (De-)Regulierung betrifft vor allem die Netzeindustrien insb. für Telekommunikation, Strom, Gas, Eisenbahn und Wasser. Ein zentrales Merkmal ihres Produktionsprozesses liegt darin, dass Investitionen für den Auf- und Ausbau eines Netzes mit meist langer Nutzungsdauer getätigt werden müssen. Da dessen Nutzung mit einem demgegenüber geringen Gütereinsatz verbunden ist, bilden die Sachkosten für Nutzung und Erhalt der Anlagen sowie die Zinsen auf das in ihnen gebundene Kapital einen hohen, häufig den wesentlichen Kostenanteil. Wenn man Kosten für diese Nutzung, beispielsweise also für ein Telefonat oder die kurzfristige Nutzung des Netzes zum Durchfluss von Informationen, Strom, Gas oder Wasser bestimmen soll, liegen daher die Kernprobleme in der Bestimmung der Abschreibungen und Zinsen (vgl. Swoboda, 1996, Küpper, 2002, S. 43 ff.). Für die Begründung kostenorientierter Entgelte der in diesen Bereichen tätigen Unternehmungen erhielten die aus dem Rechnungswesen bekannten Verfahren neue Beachtung.

In dieser Situation hatten die deutschen Unternehmungen der Strom- und Gasindustrie zuerst versucht, durch den Abschluss von „Verbände-

vereinbarungen“ freiwillig eine gegenseitige Koordination zu erreichen und dadurch die Einrichtung einer staatlichen Regulierung zu vermeiden. Diese Vereinbarungen³ waren in hohem Maße durch traditionelle Denkweisen und Verfahren der Kostenrechnung geprägt. Ausgehend von Wiederbeschaffungspreisen sollten die periodischen Abschreibungen mit den bekannten Abschreibungsverfahren ermittelt werden. Da in der Realität bei der Finanzierung eine Mischung aus Eigen- und Fremdkapital vorliegt, hatte sich durchgesetzt, dass sich entsprechend einem Prinzip der „Nettosubstanzerhaltung“ Wiederbeschaffungswerte nur auf den mit Eigenkapital finanzierten Teil beziehen könnten, während man bei dem mit Fremdkapital finanzierten Teil von den Anschaffungswerten ausgehen sollte (Busse von Colbe, 2002). Im Fall von Preisänderungen habe das jedoch zur Konsequenz, dass die Zinsen auf den eigenkapitalfinanzierten Anteil zum (für die jeweiligen Anlagegüter) ‚spezifischen‘ Realzins anzusetzen seien.

Um sich mit dieser Problematik intensiv auseinanderzusetzen, hatte die Deutsche Telekom AG Günter Knieps und mich gebeten, diesen Prozeß bei ihr gemeinsam mit Vertretern von McKinsey, insbesondere Herrn René Langen (vgl. Knieps et al., 2001), zu begleiten. Die Zusammenarbeit mit ihnen und die Diskussion insbesondere (vgl. ferner Crew & Kleindorfer, 1992) mit Stefan Reichelstein (Stanford) haben mich zu einer Reihe neuer Einsichten gebracht. Auch in unseren Lehrbüchern zur Kosten- und Erlösrechnung hatten wir bis dahin nicht ausreichend auf folgende Grundsätze hingewiesen:

1. Zwischen Abschreibungen, anderen Kosten für die Anlagennutzung wie Wartung und Instandhaltung sowie den Zinsen für das in ihnen gebundene Kapital besteht ein enger Zusammenhang. Daher sollten diese Kosten gemeinsam betrachtet werden.
2. Da es um den Einsatz längerfristig genützter Anlagen geht, ist eine ökonomische Konzeption entsprechend der Investitionstheorie und -rechnung zugrunde zu legen.
3. Durch die zu ermittelnden Kosten soll ein Gütereinsatz bzw. -verbrauch wertmäßig erfasst werden. Deshalb kann er keine Gewinnbestandteile umfassen und sollte die Abschreibungsermittlung einem Prinzip der kapitaltheoretischen Erfolgsneutralität unterliegen⁴. Bei längerfristigen Entscheidungen bezeichnet sie „den Punkt, bei dem der

3 S. z.B. VV Erdgas II (BDI et al., 2002).

4 Dies hat Busse von Colbe (2001, S. 55) nicht erkannt.

Barwert der Einzahlungen gleich dem Barwert der Auszahlungen und der Kapitalwert damit gleich null ist“ (Küpper, 2009a, S. 691).

4. Alternative Ansätze zur Bestimmung der Abschreibungen für eine Anlage

4.1 Determinanten für die Wahl des Abschreibungsverfahrens

Die Zusammenarbeit mit Günter Knieps hat mich darin bestärkt, wie wichtig theoretische Ansätze der Ökonomie auch für die Lösung konkreter praktischer Probleme sind. Besonders interessant war in der aktuellen Situation, dass die Abschreibungsermittlung nicht wie bisher üblich bei konstanten und steigenden Preisen zu untersuchen ist. Insbesondere die Unternehmungen der Telekommunikation müssen sich damit auseinandersetzen, dass die technische Entwicklung der Digitalisierung bei vielen der von ihnen eingesetzten Anlagen zu Preissenkungen führt. So wurde das Problem bei uns (vgl. aber Crew & Kleindorfer, 1992) bisher nicht gesehen und behandelt.

Die Notwendigkeit einer Berücksichtigung von Preisänderungen ergibt sich bei der Begründung von Entgelten daraus, dass die Sicht des Investors in das Netzunternehmen eingenommen werden muss (vgl. ferner Friedl, 2007). Wenn der Staat nicht die Lasten übernehmen will, sind die Preise so festzulegen, dass es Eigen- und Fremdkapitalgeber mit der Bereitschaft gibt, ihr Geld in diesen Unternehmungen anzulegen. Insofern spielt die Relevanz der Information für deren Entscheidung eine zentrale Rolle, was wir in einem Prinzip der „Entscheidungsrelevanz“ zum Ausdruck gebracht haben (vgl. auch Kretschmer et al., 2011).

Das bedeutet aber auch, dass sie nicht ohne Berücksichtigung der (künftigen) Preis-Entwicklung auf dem Beschaffungsmarkt erfolgen kann. „Wenn bei der Planung von Abschreibungen Marktentwicklungen außer Acht bleiben, besteht die Gefahr von Fehlentscheidungen und Fehlsteuerungen. Daraus leitet sich das Prinzip des *Marktbezugs* ab“ (Knieps et al., 2001, S. 760). Die damit verbundene Unsicherheit über die künftigen Entwicklungen ist in einem Risikozuschlag auf die Zinsen zu berücksichtigen (vgl. generell Pedell, 2006).

4.2 Abschreibung von Anschaffungs-, Wiederbeschaffungs- bzw. Tages- neu- und Tagesgebrauchtwerten

Traditionell werden im internen wie im externen Rechnungswesen neben der Leistungsabschreibung als wichtigste zeitabhängige Abschreibungsverfahren lineare, geometrisch- sowie arithmetisch-degressive und Annuitäten-Verfahren unterschieden (Friedl & Küpper, 2011, S. 101ff.; Küpper, 2009a, S. 694ff.). Dabei geht man in der externen Rechnungslegung von historischen Anschaffungswerten (AW) aus, bei denen der Anschaffungswert A_0 (bzw. die Differenz zwischen Anschaffungswert und Liquidationserlös LT) durch die Nutzungsdauer T dividiert wird:

$$(1) \quad a_t^{AW} = A_0 \cdot 1/T$$

In der Kostenrechnung legt man dagegen traditionell Wiederbeschaffungs- bzw. Tagesneuwerte (TNW) zugrunde. In Zusammenhang mit der Bestimmung kostenorientierter Entgelte haben Günter Knieps, René Langen und ich vorgeschlagen, stattdessen auf eine ökonomische Abschreibung überzugehen, die sich an den Tagesgebrauchtwerten (TGW) orientiert (Knieps et al., 2001, S. 764). Dabei wird in jeder Periode die Differenz zwischen den Tagesgebrauchtwerten abgeschrieben. Diese ergeben sich durch Multiplikation der Tagesneuwerte mit dem Verhältnis zwischen der Rest- und der Gesamtnutzungsdauer am Periodenanfang bzw. am Periodenende. Bezeichnet man die betrachtete Nutzungsperiode mit t sowie den Tagesneuwert für eine nicht gebrauchte Anlage am Ende dieser Periode mit A_t und wird (aus Vereinfachungsgründen) ein linearer Verschleiß unterstellt, so erhält man die TGW-Abschreibung a_t^{TGW} für die Periode t zu:

$$(2) \quad a_t^{TGW} = A_{t-1} \cdot (T-(t-1))/T - A_t \cdot (T-t)/T$$

Die Periodenkosten aus Abschreibungen und Zinsen lassen sich über Annuitäten glätten. Die Annuitätsabschreibung a_t^{Ann} ergibt sich mit dem Zinssatz i und $q=1+i$ aus der Beziehung:

$$(3) \quad a_t^{Ann} = A_0 \cdot q^T \cdot (q-1)/(q^T-1)$$

Unter Berücksichtigung einer *Preisänderungsrate* p und $\alpha=1+p$ gelangt man zu der veränderlichen Annuität:

$$(4) \quad a_t^{Ann \, P\ddot{A}} = A_0 \cdot \alpha^{t-1} \cdot (i-p) / (1-(\alpha/q)^T) = A_0 \cdot (1+p)^{t-1} \cdot (i-p) / (1-((1+p)/(1+i))^T)$$

In dem Beispiel von Abbildung 1 sind diese Verfahren für einen einzigen Investitionszyklus bei Preissenkungen von 5 % und einem nominalen

Zinssatz von 10 % berechnet. Es lässt erkennen, dass für den Ansatz von Nominalzinsen bei AW-, TGW- und Ann-Abschreibung sowie spezifischen Realzinsen bei TNW-Abschreibung alle Verfahren die Anforderung kapitaltheoretischer Erfolgsneutralität erfüllen. Zudem kann man zeigen, dass eine Trennung von Eigen- und Fremdkapital mit unterschiedlichen Zinssätzen keine qualitativ anderen Erkenntnisse liefert (Küpper, 2008, S. 79ff.).

Abbildung 1: Beispiel für verschiedene Abschreibungsverfahren bei einem Investitionszyklus

Anschaffungswert	8.000					
Preisänderungsrate $p =$	-5,00%			95,00%		
Nominalzins $i =$	10,00% $(1+i) =$			110,00%		
Realzins	15,79% $1/(1+i) =$			0,91		
Nutzungsdauer	5					
AW-Abschreibung						
Nutzungsperiode Zeitpunkt	Tages- neuwert	Abschrei- bung	Geb. EK zum Perio- denbeginn	Zinsen nominal	Perioden- summe / Entgelt	Endwert
0	8.000					
1	7.600	1.600	8.000	800	2.400	3.514
2	7.220	1.600	6.400	640	2.240	2.981
3	6.859	1.600	4.800	480	2.080	2.517
4	6.516	1.600	3.200	320	1.920	2.112
5	6.190	1.600	1.600	160	1.760	1.760
Summen	8.000					12.884
					Barwert	8.000
TNW-Abschreibung						
Nutzungsperiode Zeitpunkt	Tages- neuwert	Abschrei- bung	Geb. Kap. zum Perio- denbeginn	Zinsen real	Perioden- summe / Entgelt	Endwert
0	8.000					
1	7.600	1.520	7.600	1.200	2.720	3.982
2	7.220	1.444	5.776	912	2.356	3.136
3	6.859	1.372	4.115	650	2.022	2.446
4	6.516	1.303	2.606	412	1.715	1.886
5	6.190	1.238	1.238	195	1.434	1.434
Summen	6.877					12.884
					Barwert	8.000

TGW-Abschreibung						
Nutzungsperiode Zeitpunkt	Tages- gebraucht- wert	Abschrei- bung	Geb. Kap. zum Perio- denbeginn	Zinsen nominal	Perioden- summe / Entgelt	Endwert
0	8.000					
1	6.080	1.920	8.000	800	2.720	3.982
2	4.332	1.748	6.080	608	2.356	3.136
3	2.744	1.588	4.332	433	2.022	2.446
4	1.303	1.440	2.744	274	1.715	1.886
5	0	1.303	1.303	130	1.434	1.434
Summen		8.000				12.884
Barwert						8.000
Ann-Abschreibung						
Nutzungsperiode Zeitpunkt	Tages- neuwert	Abschrei- bung	Geb. EK zum Perio- denbeginn	Zinsen nominal	Perioden- summe / Entgelt	Endwert
0	8.000					
1	7.600	1.510	8.000	800	2.310	3.382
2	7.220	1.545	6.490	649	2.194	2.921
3	6.859	1.590	4.945	495	2.085	2.522
4	6.516	1.645	3.355	336	1.980	2.178
5	6.190	1.710	1.710	171	1.881	1.881
Summen		8.000				12.884
Barwert						8.000

Quelle: Eigene Darstellung

Die handelsrechtliche Bedingung, dass die Summe der Abschreibungsbe-
träge den historischen Anschaffungswert ergibt, wird verletzt, wenn man
von den Tagesneuwerten als Wiederbeschaffungswerten ausgeht. Mit den
jeweiligen Tagesgebrauchtwerten wird diese Bedingung dagegen eingehal-
ten. Daher lässt sich argumentieren, dass die ökonomische⁵ oder TGW⁶-
Abschreibung der linearen und der TNW-Abschreibung bei Preisänderun-
gen überlegen ist. Deshalb wird im Folgenden die TNW-Abschreibung
nicht mehr in den Vergleich einbezogen.

5 So haben wir sie aufgrund ihres Bezugs zur Ertragswertabschreibung in Knieps et
al. (2001) benannt.

6 Wegen des deutlicheren Bezugs zum zugrundeliegenden Wert in den nachfolgen-
den Aufsätzen.

Die Verwendbarkeit der TGW-Abschreibung für Entscheidungszwecke erschien uns „in den Fällen höher, wo die Unternehmung damit rechnen muss, dass auf Grund von Marktzusammenhängen oder Regulierungsbedingungen eine Beziehung zwischen den Preisentwicklungen auf ihren Beschaffungs- und Absatzmärkten besteht“ (Knieps et al., 2001, S. 767). Sind keine zuverlässigen Annahmen über künftige Entwicklungen möglich, müssen die Anlagen entsprechend dem Konzept „offener“ Abschreibungspläne periodisch neu bewertet werden. Dann ergibt sich „der ökonomische Wertverzehr ... jeweils als Differenz zwischen dem bekannten Wert der Anlage zu Beginn der Periode und dem (erwarteten) Wert der Anlage am Ende der Periode“ (Knieps et al., 2001, S. 769), wie es die ökonomische oder TGW-Abschreibung vorsieht. Aus diesen Gründen hatten wir damals der Deutschen Telekom AG empfohlen, von der Abschreibung von Anschaffungswerten auf dieses Verfahren überzugehen.

4.3 Tagesgebrauchtwert- oder Annuitätenabschreibung?

Eine interessante Erweiterung erhielt die Analyse durch die Fundierung der Annuitätenabschreibung über einen Ansatz, der auf Arrow (1964) zurückgeht und von Rogerson (2008, 2011), Nezlobin et al. (2012) sowie Rajan und Reichelstein (2009) vorgeschlagen wurde. Sie leiten den Wettbewerbspreis einer Kapazitätseinheit (competitive rental price of capacity (Nezlobin et al., 2012)) aus einem Modell ab, in dem die effiziente (=optimale) Anlagenkapazität durch Maximierung der Konsumentenwohlfahrt bestimmt wird. Dies erfüllt (für einen gegebenen Zinssatz) die Bedingung kapitaltheoretischer Erfolgsneutralität⁷ und führt bei konstanten Preisen sowie gleichbleibender Kapazitätsnutzung zu annuitätisch konstanten Periodenkosten:

$$(6) \quad a_t^{\text{RRC}} = A_0 \cdot 1 / \sum_{j=1}^T (1+i)^{-j}$$

Das sich ergebende Kostenverteilungs- oder Abschreibungsverfahren wird von Rogerson als „Relative Replacement Cost“- (RRC-) Regel bezeichnet, die bei Existenz eines Wettbewerbsmarkts für Anlagen mit Hotellings (1925) Konzept der ökonomischen Abschreibung übereinstimmt (Nezlobin et al., 2012, S. 9, Fn. 18). Man kann zeigen, dass diese Annuität mit der von Gleichung 3 und im Fall von *Preisänderungen* mit der Rate p mit derje-

7 Nezlobin et al. (2012) sprechen von einem „constraint that the firm does not receive direct monetary transfers from the regulator“.

nigen von Gleichung 4 übereinstimmt (vgl. Friedl & Küpper, 2011, S. 126):

$$(7) \quad a_t^{\text{RRC-P}\ddot{A}} = A_0 \cdot (1+p)^t \cdot 1 / \sum_{j=1}^T (1+p)^j \cdot (1+i)^{-j}$$

Die TGW- und die Ann-Abschreibung erscheinen ökonomisch gut begründet. Offensichtlich führen beide Verfahren nicht zu denselben Periodenkosten. Damit stellt sich die Frage, worin ihre Unterschiede liegen, worauf diese zurückzuführen sind und welches Verfahren sich besser begründen lässt.

Um ihre strukturellen Differenzen herauszufinden, ist der einfachste denkbare Fall einer Anlage mit einer Nutzungsdauer von 2 Perioden und konstanten (Anschaffungs-)Preisen zu untersuchen. Bei diesem geht die TGW- in die AW-Abschreibung über. Das in Abbildung 2 wiedergegebene Beispiel für eine Anlage mit einem Anschaffungswert $A=8.000$ und Nominalzinsen von $i=10\%$ zeigt, dass sich die Periodenkosten trotz gleichem Barwert sowohl je Periode als auch in der Summe unterscheiden.

Abbildung 2: AW- bzw. TGW- und Ann-Abschreibung bei 2 Investitionszyklen ohne Preisänderungen

Anschaffungswert		8.000			
Nominalzins		10,00%			
Nutzungsdauer		2			
AW-Abschreibung		Perioden-			
t	BW	Abschr	Zinsen	kosten	Barwert
0	8.000				
1	4.000	4.000	800	4.800	4.364
2		4.000	400	4.400	3.636
		8.000	1.200	9.200	8.000
Ann.-Abschreibung		Perioden-			
t	BW	Abschr	Zinsen	kosten	Barwert
0	8.000				
1	4.190	3.810	800	4.610	4.190
2		4.190	419	4.610	3.810
		8.000	1.219	9.219	8.000

Quelle: Eigene Darstellung

Die Ann-Abschreibung hat wegen der zwangsläufig höheren Zinsen in der ersten Periode eine geringere und in der zweiten Periode eine höhere Ab-

schreibung. Dies ist die unvermeidliche Folge des Tatbestands, dass die Periodenkosten in einem Annuitätenansatz sowie nach dem auf Arrow zurückgehenden Konzept bei unveränderlichen Anlagenpreisen konstant sind. Für Anlagen mit einer Nutzungsdauer von 2 Perioden lässt sich dies aus dem entscheidungstheoretischen Ansatz leicht nachvollziehbar herleiten (vgl. hierzu Nezlobin et al., 2012). In dem hierbei formulierten Optimierungsmodell stellen die Periodenkosten jene langfristigen Zusatzkosten dar, die durch die Erhöhung der Kapazität in der gegenwärtigen Periode um eine Einheit ausgelöst werden. Diese entsprechen dem Barwert der Änderungen in der Anschaffung der Anlagen in den nachfolgenden Perioden (vgl. hierzu ausführlicher Friedl & Küpper, 2011, S. 108f.; Küpper, 2009a, S. 704ff.). Geht man davon aus, dass die Unternehmung einen Bestand derartiger Anlagen mit einer Nutzungsdauer $T=2$ aufbaut, führt der Kauf einer Anlage in $t=0$ dazu, dass in $t=1$ auf den ursprünglich geplanten Kauf einer Anlage verzichtet werden kann. Damit fehlt deren Kapazität in der nachfolgenden Periode, so dass in $t=2$ eine weitere Anlage zu kaufen ist. Diese Verschiebung der Anschaffungszeitpunkte setzt sich in einer unendlichen Reihe fort.

Das zugrunde liegende entscheidungstheoretische Konzept betrachtet aber nicht eine einzelne Anlage, sondern mindestens die Zahl an Anlagen, die der Nutzungsdauer entspricht. Nur in diesem Fall führt die Zusatzinvestition zum Betrachtungszeitpunkt $t=0$ zu den Veränderungen in den Folgeperioden, aus denen die langfristigen Zusatzkosten hergeleitet werden. Die Periodenkosten bei Annuitätenabschreibung liegen ab der Periode mit einem ausgeglichenen Anlagenbestand um die Differenz der Kosten je Periode zwischen AW- und Ann-Abschreibung während des Kapazitätsaufbaus systematisch über denjenigen der AW-Abschreibung. Die Differenz zwischen TGW- (bzw. ohne Preisänderungen AW-) und Annuitätenabschreibungen ohne und mit Preisänderungen ist demnach darauf zurückzuführen, dass die TGW-Abschreibung die Wertentwicklung der einzelnen Anlagen beim Aufbau eines Anlagenbestands explizit einbezieht, während die Ann-Abschreibung die durchschnittlich gleichen Periodenkosten auch schon für die erste Periode ansetzt. Da die Abschreibungen einer Anlage und demzufolge auch die Periodenkosten beim TGW-Verfahren in den ersten Perioden stets höher als beim Annuitätenverfahren sind, liegen dafür die Periodenkosten bei Annuitäten nachfolgend für den (ausgeglichenen) Anlagenbestand stets über denjenigen des TGW-Verfahrens.

Dem TGW-Verfahren liegt die Hypothese zugrunde, dass sich Abschreibungen aus den Differenzen zwischen den Gebrauchtwerten für eine Anlage in den jeweiligen Perioden ergeben. Sie wirkt ökonomisch unmittelbar

einleuchtend. Wenn auf dem Markt Gebrauchtwerte existieren, zu denen Anlagen mit denselben technologischen Eigenschaften ge- und verkauft werden, ist dies der anzusetzende Wert. Daher stellt sich die Frage, ob und inwieweit diese Werte durch Gleichung 2 korrekt wiedergegeben werden. Dieser berücksichtigt zwar die Preisänderung, jedoch keine Zeitpräferenzen und geht für konstante Preise in die AW-Abschreibung über. Darüber hinaus bezieht er die Zinsen für das in der Anlage gebundene Kapital nicht ein. Aus ökonomischer Sicht ist jedoch davon auszugehen, dass Investoren bei ihren Entscheidungen auch die Zinsen beachten und sich diese deshalb in den Marktpreisen niederschlagen müssten. Dann gibt Gleichung 2 die Bildung von Tagesgebrauchtwerten unvollständig wieder und das Annuitätenverfahren erscheint als ökonomisch besser begründet.

5. Bedeutung der Unterschiede zwischen den Verfahren bei ausgeglichenem Anlagenbestand

Die Analyse in Abschnitt 4.2 betrachtet die Abschreibungen, Zinsen und Periodenkosten von AW-, TGW- und Ann-Abschreibung für eine Anlage. Der Vergleich der Periodenkosten des Beispiels von Abbildung 1 deutet darauf hin, dass ihre Differenzen nicht vernachlässigt werden sollten. Die Unterschiede zwischen der in der Praxis bevorzugten AW-Abschreibung und den Verfahren, welche Preisänderungen berücksichtigen, betragen in diesem Beispiel bis zu mehr als 20 % (vgl. Abb. 3).

Abbildung 3: Verhältnis zwischen den Periodenkosten bei einmaligem Investitionszyklus

Abschreibungs- Verfahren	Kosten in Periode				
	1	2	3	4	5
AW-Abschreibung	2.400	2.240	2.080	1.920	1.760
TNW-Abschreibung	2.720	2.356	2.022	1.715	1.434
TGW-Abschreibung	2.720	2.356	2.022	1.715	1.434
Ann-Abschreibung	2.310	2.194	2.085	1.980	1.881
	Verhältnis der Kosten in Periode:				
	1	2	3	4	5
AW-/TNW-Abschr.	0,88	0,95	1,03	1,12	1,23
AW-/TGW-Abschr.	0,88	0,95	1,03	1,12	1,23
AW-/Ann-Abschr.	1,04	1,02	1,00	0,97	0,94
Ann-/TGW-Abschr.	0,85	0,93	1,03	1,15	1,31

Quelle: Eigene Darstellung

Eine Erweiterung der Sichtweise und der Beurteilung ergibt sich aber, wenn man sie auf einen – möglichst ausgeglichenen – Anlagenbestand erweitert. Schon in dem auf Arrow zurückgehenden Optimierungsansatz wird nicht eine einzelne Anlage, sondern eine unendliche Kette von Investitionen unterstellt (Arrow, 1964; Nezlobin et al., 2012). In einem „eingeschwungenen“ Zustand ist der Anlagenbestand insoweit ‚ausgeglichen‘, dass er mindestens so viele Anlagen wie deren maximale Nutzungsdauer umfasst und laufende Reinvestitionen erfolgen. In Abbildung 4 werden dazu beispielhaft die Periodenkosten eines ausgeglichenen Anlagenbestands von 5 Anlagen des Beispiels von Abbildung 1 mit einer Nutzungsdauer von 5 bei $i=10\%$ verglichen. Dabei zeigt sich, dass die Abweichungen zwischen den Verfahren in allen Perioden gleich und mit bis zu ca. 3 % deutlich geringer als für einen einzigen Investitionszyklus sind.

Abbildung 4: Verhältnis der Periodenkosten bei ausgeglichenem Anlagenbestand mit 5 Anlagen einer Nutzungsdauer von 5 des Beispiels von Abbildung 1

Abschreibungs- Verfahren	Kosten in Periode:				
	1	2	3	4	5
AW-Abschreibung	11.463	10.890	10.345	9.828	9.337
TGW-Abschreibung	11.200	10.640	10.108	9.603	9.122
Ann-Abschreibung	11.549	10.971	10.423	9.901	9.406
	Verhältnis der Kosten in Periode: 1				
	2	3	4	5	
AW-/TGW-Abschr.	1,023	1,023	1,023	1,023	1,024
AW-/Ann-Abschr.	0,993	0,993	0,993	0,993	0,993
Ann-/TGW-Abschr.	1,031	1,031	1,031	1,031	1,031

Quelle: Eigene Darstellung

Die Berücksichtigung des Tatbestands, dass die Unternehmungen der Netzindustrie in der Realität ihre Leistungen mit einer Vielzahl von Anlagen erbringen, die relativ regelmäßig ersetzt werden, führt somit zu einer neuen Beurteilung. Die Unterschiede erweisen sich als weniger dramatisch, als es der Vergleich für Einzelanlagen erscheinen lässt. Sie fallen um so weniger ins Gewicht, je regelmäßiger Reinvestitionen durchgeführt werden und je kürzer die Nutzungsdauern sind.

Wie aber die Ergebnisse einer Simulation unterschiedlicher Kombinationen von Nutzungsdauern, Zinssätzen und Preisänderungsraten (vgl. hierzu ausführlich Friedl & Küpper, 2011, S. 109ff.; Küpper & Pedell,

2016, S. 95ff.) in Abbildung 5⁸ veranschaulichen, gilt dies jedoch nur, wenn sich Zinsen, Preisänderungen und Nutzungsdauern in unteren Bereichen bewegen. Zu deutlichen und gravierenden Differenzen insbesondere zum AW-Verfahren kommt es vor allem bei langen Nutzungsdauern.

Abbildung 5: Beispiele für das Verhältnis der Periodenkosten von AW-, TGW- und Ann-Abschreibung bei unterschiedlichen Zinssätzen, Preisänderungsraten und Nutzungsdauern

i	p	T	AW/Ann	TGW/Ann	i	p	T	AW/Ann	TGW/Ann
5%	-20 %	5	1,101	0,922	16,6%	-20%	5	1,084	0,879
		10	1,478	0,813			10	1,447	0,751
		20	4,228	0,682			20	4,194	0,634
5%	-10 %	5	1,019	0,967	16,6%	-10%	5	0,997	0,927
		10	1,086	0,904			10	1,035	0,822
		20	1,408	0,787			20	1,337	0,690
5%	0%	5	0,996	0,996	16,6%	0%	5	0,967	0,967
		10	0,985	0,985			10	0,904	0,904
		20	0,950	0,950			20	0,788	0,788
5%	10%	5	1,013	0,995	16,6%	10%	5	0,977	0,994
		10	1,052	0,997			10	0,923	0,979
		20	1,204	0,883			20	0,791	0,935
5%	20%	5	1,061	0,950	16,6%	20%	5	1,016	0,998
		10	1,265	0,700			10	1,059	0,992
		20	2,327	-1,681			20	1,195	0,963

Quelle: Eigene Darstellung

6. Ergebnis: Lessons learned? – Eine Aufgabe der Umsetzung!

Der von der Regulierung ausgehende Anstoß zur erneuten Befassung mit dem Abschreibungsproblem hat zu wichtigen Erkenntnissen geführt. Damals ist es Günter Knieps, Robert Langen und mir nicht gelungen, die Telekom AG zu einem Übergang auf ein neueres Abschreibungsverfahren zu überzeugen. Wie die hier vorgenommenen Vergleiche zeigen, erscheint das Festhalten an der handels- und steuerrechtlich gut abgesicherten Anschaffungswertabschreibung gerechtfertigt, solange sich die Zinsen, Preisänderungen und Nutzungsdauern im begrenzten Rahmen bewegen.

Diese Voraussetzung ist in Netzindustrien im Hinblick auf für viele Anlagen doch recht lange Nutzungsdauern nicht erfüllt. Deshalb sollten die

8 Die Werte sind ermittelt aus Friedl & Küpper (2011, S. 111f.).

damals auch international erarbeiteten Erkenntnisse nicht nur in der wissenschaftlichen Diskussion berücksichtigt werden und in die Lehrbücher (vgl. z.B. Schweitzer et al., 2016, S. 124ff.) eingehen. Sie sollten auch von den betroffenen Unternehmungen aufgegriffen werden. Das führt zurück zu der Diskussion um die ökonomische bzw. Tagesgebrauchtwert- oder die Annuitätenabschreibung. Letztere ist theoretisch besser fundiert, wenn man die Tagesgebrauchtwerte lediglich von der vermuteten Preisänderungsrate abhängig macht. Kann man dagegen auf empirisch beobachtbare Tagesgebrauchtwerte und begründete Hypothesen über deren Entwicklung zurückgreifen, sprechen mehrere Gründe für deren Verwendung. Einmal stimmt bei ihr (wie bei der Annuitätenabschreibung) die Summe der Abschreibungen mit dem Anschaffungswert überein. Hierdurch könnte sie auch in der externen Rechnungslegung anwendbar sein, zumal man zu ihrer Bestimmung jeweils nur den beobachtbaren (beizulegenden) Tageswert am Anfang und den zu prognostizierenden Wert am Ende dieser (nächsten) Periode benötigt, ist sie „offen“. Der theoretischen Begründung für die Annuitätenabschreibung liegt dagegen eine längerfristige Sicht zugrunde. Die Tageswertabschreibung ist daher ‚robuster‘ gegenüber Preisänderungen. Zudem liegt sie konzeptionell näher an der Anschaffungswertabschreibung als die Annuitätenabschreibung. Aus diesen Gründen erscheint mir ihre Umsetzung in der Praxis leichter erreichbar.

Die erneute Analyse des Abschreibungsproblems deutet auf einen grundlegenden Aspekt hin, der sich bei ihrer Bestimmung für unternehmerische Entscheidungsprobleme noch klarer gezeigt hat.⁹ Maßgeblich für ihre Bestimmung sind weniger vergangene als die gegenwärtigen und künftigen Anlagenwerte. Soweit Abschreibungen als Informationen für (externe) Entscheidungen in der Regulierung über künftige Preise ihrer Nutzung oder für innerbetriebliche Entscheidungen über ihre Nutzung bereitzustellen sind, hat sich der Blick grundsätzlich zu ändern. Ihre Bestimmung sollte weniger auf vergangene Anschaffungs- als auf zukünftige Gebrauchtwert- und Anschaffungswerte blicken.

9 Vgl. hierzu die Ansätze zur Bestimmung nutzungsabhängiger Abschreibungen für Entscheidungsprobleme mithilfe des investitionstheoretischen Ansatzes (Küpper, 1985, 2009c).

Referenzen

- Arrow, K. J. (1964). Optimal capital policy, the cost of capital, and myopic decision rules. *Annals of the Institute of Statistical Mathematics*, 16(1), 21–30. <https://doi.org/10.1007/BF02868559>.
- BDI, VIK, BGW, & VKU. (2002). *Verbändevereinbarung zum Netzzugang bei Erdgas vom 3. Mai 2002 (VV Erdgas II)*. Bundesverband der Deutschen Industrie e.V. (BDI), Verband der Industriellen Energie- und Kraftwirtschaft e.V. (VIK), Bundesverband der deutschen Gas- und Wasserwirtschaft e.V. (BGW) & Verband kommunaler Unternehmen e.V. (VKU).
- Busse von Colbe, W. (2001). Kostenorientierte Entgeltregulierung von Telekommunikationsdienstleistungen bei sinkenden Beschaffungspreisen für Investitionen. In U. Wagner (Ed.). *Zum Erkenntnisstand der Betriebswirtschaftslehre am Beginn des 21. Jahrhunderts: Festschrift für Erich Loitsberger zum 80. Geburtstag* (S. 47–59). Berlin: Duncker & Humblot.
- Busse von Colbe, W. (2002). Zur Ermittlung der Kapitalkosten als Bestandteil regulierter Entgelte für Telekommunikationsdienstleistungen. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 54(S48), 1–25.
- Crew, M. A., & Kleindorfer, P. R. (1992). Economic Depreciation and the Regulated Firm under Competition and Technological Change. *Journal of Regulatory Economics*, 4(1), 51–61.
- Friedl, G. (2007). Ursachen und Lösung des Unterinvestitionsproblems bei einer kostenbasierten Preisregulierung. *Die Betriebswirtschaft*, 67(3), 335–349.
- Friedl, G., & Küpper, H.-U. (2011). Historische Kosten oder Long Run Incremental Costs als Kostenmaßstab für die Preisgestaltung in regulierten Märkten?. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 63(S64), 98–128. <https://doi.org/10.1007/BF03373003>.
- Hotelling, H. (1925). A General Mathematical Theory of Depreciation. *Journal of the American Statistical Association*, 20(151), 340–353.
- Knieps, G. (2007). *Netzökonomie: Grundlagen, Strategien, Wettbewerbspolitik*. Wiesbaden: Gabler Verlag.
- Knieps, G., Küpper, H.-U., & Langen, R. (2001). Abschreibungen bei fallenden Wiederbeschaffungspreisen in stationären und nicht stationären Märkten. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 53(8), 759–776. <https://doi.org/10.1007/BF03372667>.
- Kosiol, E. (1964). *Kostenrechnung*. Wiesbaden: Gabler Verlag.
- Kretschmer, T., Küpper, H.-U., & Pedell, B. (2011). Bedingungen und Prinzipien einer konsistenten Regulierung. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 63(S64), 1–31. <https://doi.org/10.1007/BF03372999>.
- Küpper, H.-U. (1984). Kosten- und entscheidungstheoretische Ansatzpunkte zur Behandlung des Fixkostenproblems in der Kostenrechnung. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 36(10), 794–811.
- Küpper, H.-U. (1985). Investitionstheoretische Fundierung der Kostenrechnung. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 37(1), 26–46.

- Küpper, H.-U. (2002) Kostenorientierte Preisbestimmung für regulierte Märkte – Analyse eines Beispiels der Bedeutung betriebswirtschaftlicher Begriffe und Konzepte. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 48 (S02), 27–55.
- Küpper, H.-U. (2008). Marktorientierte kapitaltheoretische Bestimmung kostenorientierter Entgelte für regulierte Unternehmungen. In O. Altenburger (Ed.). *Steuern: Steuern, Steuerung, Regulierung und Prüfung* (S. 67–92). Wien: Linde Verlag.
- Küpper, H.-U. (2009a). Anschaffungs- oder Wiederbeschaffungs- bzw. Tageswerte bei Abschreibungen? In T. Egner, K. Henselmann, & L. Schmidt (Eds.). *Steuern und Rechnungslegung: Festschrift zum 65. Geburtstag von Professor Dr. Jochen Sigloch* (S. 687–713). Düren: Shaker Verlag.
- Küpper, H.-U. (2009b). Effizienzreform der deutschen Hochschulen nach 1990 – Hintergründe, Ziele, Komponenten. *Beiträge Zur Hochschulforschung*, 31(4), 50–75.
- Küpper, H.-U. (2009c). Investment-Based Cost Accounting as a Fundamental Basis of Decision Oriented Management Accounting. *Abacus* 45 (2), 249–274.
- Küpper, H.-U., & Pedell, B. (2016). Which asset valuation and depreciation method should be used for regulated utilities? An analytical and simulation-based comparison. *Utilities Policy*, 40, 88–103. <https://doi.org/10.1016/j.jup.2016.05.001>.
- Mellerowicz, K. (1933). *Kosten und Kostenrechnung. I. Theorie der Kosten, II. Kostenrechnung*. Berlin und Leipzig: Walter de Gruyter
- Nezlobin, A., Rajan, M. V., & Reichelstein, S. (2012). Dynamics of Rate-of-Return Regulation. *Management Science*, 58(5), 980–995. <http://www.jstor.org/stable/41499533>.
- Pedell, B. (2006). *Regulatory Risk and the Cost of Capital: Determinants and Implications for Rate Regulation*. Berlin & Heidelberg: Springer.
- Rajan, M. V., & Reichelstein, S. (2009). Depreciation Rules and the Relation between Marginal and Historical Cost. *Journal of Accounting Research*, 47(3), 823–865. <https://doi.org/10.1111/j.1475-679X.2009.00334.x>.
- Rogerson, W. P. (2008). Intertemporal Cost Allocation and Investment Decisions. *Journal of Political Economy*, 116(5), 931–950. <https://doi.org/10.1086/591909>.
- Rogerson, W. P. (2011). On the relationship between historic cost, forward-looking cost and long run marginal cost. *Review of Network Economics* 10(2). Article 2.
- Schmalenbach, E. (1925). *Grundlagen der Selbstkostenrechnung und Preispolitik*. (2. Auflage). Leipzig: Glöckner.
- Schweitzer, M., Küpper, H.-U., Friedl, G., Hofmann, C., & Pedell, B. (2016). *Systeme der Kosten- und Erlösrechnung* (11., überarbeitete und erweiterte Auflage). München: Verlag Franz Vahlen.

Der Deutschland-Takt: Regulierung, Wettbewerb und Organisation auf der Schiene

Kay Mitusch¹

Abstract

Deutschland-Takt ist zunächst ein konzeptioneller Rahmen für eine fahrplanbasierten Entwicklung der Schieneninfrastruktur, d.h. ein Investitionsprogramm. Darüberhinaus sieht er eine Weiterentwicklung des Fahrplans im Schienenpersonenverkehr in Richtung auf einen integrierten Taktfahrplan vor. Dies wirft jedoch organisatorische und wettbewerbspolitische Fragen auf. Dieser Artikel zeigt, dass sowohl Wettbewerb im Deutschlandtakt als auch Wettbewerb zum Deutschlandtakt möglich und wünschbar sind. Doch warnt er davor, Verkehrsangebote im Deutschland-Takt direkt zu subventionieren, da dies schnell anreiz- und politökonomisch unbeherrschbar werden kann.

Keywords:

Eisenbahn, Schienenpersonenverkehr, SPfV, Wettbewerb im Deutschland-Takt, Wettbewerb zum Deutschland-Takt

1. Einleitung

Vielleicht ist „Deutschland-Takt“ nicht viel mehr als ein Investitionsprogramm des Bundes, um endlich die vielen kleinen und großen Engpässe im Schienennetz in Angriff zu nehmen. Vorgenommen hat man sich das schon seit langem. Das Geld, welches immer fehlte, ist seit etwa 2016 und insb. 2019 da (wie lange noch?), und man braucht einen geeigneten kon-

1 Karlsruher Institut für Technologie (KIT), Institut für Volkswirtschaftslehre (ECON), Netzwerkökonomie. mitusch@kit.edu. Auch IGES Institut GmbH, Berlin. Der Artikel beruht auf der Wissenslandkarte „Potenziale und Herausforderungen des Deutschland-Takts im Schienenpersonenfernverkehr“ des Forschungsinformationssystems (FIS, 2021). Vielen Dank an Jannik Dresselhaus für seine Beiträge für die Wissenslandkarte und diesen Artikel.

zeptionellen Rahmen, um es nicht nur für einzelne Großprojekte, sondern für die Beseitigung von Engpässen im ganzen Netzkontext vernünftig einsetzen zu können. Die für den Deutschland-Takt grundlegende Idee der „fahrplanbasierten Entwicklung der Schieneninfrastruktur“ (SMA und Partner et al., 2021, S. 88) ist hierfür genau das Richtige. Und zwar selbst dann, wenn hinterher die weitergehenden Ideen eines „integralen Taktfahrplans“ nur sehr unvollkommen in die Tat umgesetzt werden.

Mehr als „nur“ ein Infrastruktur-Investitionsprogramm ist der Deutschland-Takt dann und in dem Maße, wie der integrale Taktfahrplan im Schienenpersonenverkehr (SPV) realisiert und vom Staat organisiert und finanziert wird. Integraler Taktfahrplan (ITF) heißt im Idealfall: An jedem beteiligten Bahnhof treffen sich zur bundesweit festgelegten Taktzeit (also z.B. um 10.30 Uhr und dann jede Stunde) alle Züge, die dort fahren, und erlauben das Umsteigen in allen Kombinationen. Dann fahren sie gemeinsam los, um zur nächsten Taktzeit (oder, bei längeren Strecken, ganzzahligen Vielfachen der Taktzeit) am nächsten Bahnhof mit anderen Zügen zusammenzutreffen. Dass dies in dieser strengen Form in Deutschland nicht umsetzbar ist – oder jedenfalls lieber nicht umgesetzt werden sollte – ist bekannt. Echter Hochgeschwindigkeitsverkehr (HGV) sollte jedenfalls nicht gezwungen werden, jede Stunde zu halten und längere Zeit auf Umsteiger zu warten. Daher strebt man für Deutschland einen ITF mit vielen Kompromissen an: Mehr Takt und Umsteigefähigkeit sollen ermöglicht werden, aber nur soweit dies sinnvoll ist, nicht das vollständige Idealmodell (ARGE IGES Institut & IVE, 2015; Berschin, 2015; SMA und Partner et al., 2021).²

Bekannt ist auch, dass der aktuelle Fahrplan des Marktführers DB Fernverkehr AG (im Folgenden: „DB Fern“) von einem integralen Taktfahrplan gar nicht so weit entfernt ist. Ein Hauptmanko ist derzeit die mangelnde Verlässlichkeit bei der Durchführung des Fahrplans. Viele Optimierungen des Fahrplans zur Ermöglichung von (pünktlichen) Taktungen und Umsteigefähigkeiten werden durch Infrastrukturengpässe, insb. im Umfeld der Knoten, also der großen Bahnhöfe wie Hamburg, Köln, Frankfurt und Mannheim, verhindert. Daher kann es durchaus sein, dass nach Umsetzung des Infrastruktur-Investitionsprogramms des Bundes die DB Fern ihren Fahrplan von sich aus einem ITF weiter annähern wird, weil es dann möglich wird. Oder dass Wettbewerber dies erledigen, indem

2 Zum Deutschland-Takt allgemein siehe auch die offizielle Internetseite des BMDV: www.deutschlandtakt.de.

sie sich mit ihren Angeboten in den Fahrplan der DB Fern „integral konkurrierend“ einklinken. Dann bräuchte der Staat nichts zu organisieren.

Die verkehrspolitischen Vorstellungen gehen aber weiter: Engerer Takt, am besten halbstündlich; günstigere Fahrpreise; mehr Städte an den vertakteten Schienenpersonenfernverkehr (SPFV) anbinden; den SPFV-Takt noch enger mit dem regionalen Schienenpersonennahverkehr (SPNV) integral vertakten usw. Die Idee des Deutschland-Takts ist im Umfeld der Aufgabenträger des SPNV entstanden und gereift, und wurde von den Ländern über den Bundesrat auf die Tagesordnung der Bundespolitik gehieft (Berschlin et al., 2019, S. 6; BRat, 2016). Als Vorbilder gelten Beispiele aus den Niederlanden und der Schweiz und aus regionalen SPNV-Netzen. Der „Mindset“ der Aufgabenträger des SPNV lässt sich mit vier Begriffen umreißen: Öffentliche Organisation, massive staatliche Finanzierung, am liebsten ein Verbot von direktem Wettbewerb³, aber dafür Ausschreibungswettbewerb. Und so wird vielfach auch der Deutschland-Takt gedacht, wenn er deutlich mehr als ein Infrastruktur-Investitionsprogramm des Bundes sein soll. Daneben gibt es auch softere Ideen wie die eines „Deutschland-Tarifs“, der auf stärkere Durchbindung von Tickets zwischen SPFV und SPNV sowie Vereinheitlichung von Tarifkategorien (wie werden zum Beispiel „Kind“ und „Senior“ definiert?) dringt (Leister, 2013).

Alles was über die Infrastruktur-Investitionen hinausgeht, wirft regulatorische, ordnungspolitische und politökonomische Fragen auf. Im Kern geht es um das Spannungsfeld zwischen Wettbewerb und Regulierung. Dies wird angereichert durch die Frage der Universaldienstverpflichtung oder „Daseinsvorsorge“: Welche Mindestversorgungsqualität wird flächendeckend gefordert, und wie ist sie zu garantieren und zu finanzieren? Das sind klassische Fragen der Netzwerkökonomie, die sich in unterschiedlicher Ausprägung auch in anderen Netzwerksektoren stellen. Somit sind es auch die zentralen Fragen im wissenschaftlichen Werk von Günter Knieps, der sie sowohl sektorübergreifend als auch in Hinblick auf den Verkehr und die Eisenbahn intensiv bearbeitet hat (Knieps, 1996, 2007, 2008; Knieps & Brunekreeft, 2003; Knieps & Weiß, 2009).

3 Im SPNV gibt es zwar kein Wettbewerbsverbot, aber aufgrund der dichten, subventionierten Verkehrsangebote gab es bisher keinen konkurrierenden Markteintritt. Bei der Markttöffnung für den Fernbusverkehr wurde hingegen das Wettbewerbsverbot für den „Nah“-Verkehr beibehalten, und zwar mit der weiten 50 km-Grenze, um nicht nur die ÖPNV-Stadtverkehre, sondern auch den SPNV gegen Konkurrenz abzusichern. Paradoxerweise könnte also Flixmobility konkurrierenden Regionalverkehr auf der Schiene, nicht jedoch auf der Straße anbieten.

Diese Fragen stehen auch im Zentrum des vorliegenden Artikels. Vorab wird in Abschnitt 2 unter der Überschrift „ITF versus HGV“ die oben schon angerissene Frage nach der geeigneten, abgeschwächten Form eines ITF für Deutschland adressiert. Ein Deutschland-Takt kann mehr oder weniger stark institutionell organisiert werden. Drei verschiedene Varianten mit zunehmendem Grad der institutionellen Verankerung werden in Abschnitt 3 skizziert mit dem Fokus auf der Frage, wie im jeweiligen Rahmen „Wettbewerb im und zum Deutschland-Takt“ ermöglicht, gefördert oder sogar organisiert werden kann. Sollten die Verkehrsangebote des Deutschland-Takts finanziell gefördert werden (ähnlich zum SPNV), stellen sich darüberhinaus Fragen finanzieller, anreizökonomischer und politökonomischer Art, die in Abschnitt 4 unter der Überschrift „Eigenwirtschaftlichkeit versus Bestellerprinzip im Deutschland-Takt“ diskutiert werden. Die Ergebnisse zusammenfassend unterbreitet Abschnitt 5 „Fünf Vorschläge zur weiteren Entwicklung des Deutschland-Takts“.

2. ITF versus HGV

Der Fahrplan der DB Fern deckt kreuz und quer die gesamte Fläche des Landes ab und ist regelmäßig getaktet. Allerdings sind an vielen Kreuzungspunkten die Umsteigemöglichkeiten unzureichend, insbesondere sind sie selten so „integral“ gestaltet, dass man zwischen allen Richtungen hin- und herwechseln könnte. Vielmehr fühlt sich oft ein Teil der Umsteigepassagiere wie gestrandet, wenn man ein oder gar zwei Stunden auf die nächste Anschlussverbindung warten muss, weil die letzte kurz vorher, aber unerreichbar abgefahren ist.

In den Zeiten der Deutschen Bundesbahn sah der Fahrplan einem ITF noch ähnlicher als heute.⁴ Viele Stimmigkeiten lösten sich (auch schon zu Zeiten der Deutschen Bundesbahn) mit dem Bau der HGV-Strecken auf, vielleicht nur aus Unachtsamkeit. So ist beispielsweise die 1991 eröffnete Strecke von Mannheim nach Stuttgart mit einer Reisedauer von mindestens 36 Minuten ungeeignet für einen ITF, in welchem die Fahrzeiten zwischen zwei Knoten Vielfache einer festgelegten Taktzeit (meist 30 oder 60 Minuten) betragen sollten (Hesse, 2019).

Aber passen HGV und ITF überhaupt zusammen? Im HGV fahren die Züge mit konstant hoher Geschwindigkeit, anstatt immer wieder zu halten

4 Siehe Berschin (2015, S. 635) und den Bericht „Historie des Taktfahrplans im deutschen SPNV“ in FIS (2021).

und auf Umsteiger zu warten. So wurden die HGV-Verbindungen oftmals als Expresslinien geplant, die keineswegs überall in den Fahrplan der anderen Linien (damals IC- und D-Züge) eingeflochten sein müssen (ARGE IGES Institut & IVE, 2015, S. 8f.). Dann brachte man es aber doch nicht über's Herz, die HGV-Strecken an größeren Städten vorbeizuführen oder die Züge dort einfach durchfahren zu lassen. So entwickelte sich angesichts der vielen mittelgroßen Städte in Deutschland eine Kompromisslösung: recht schnelle, aber nicht sehr schnelle Verbindungen über die großen Distanzen.

Wie sieht es im Ausland aus? In der kleinen Schweiz, dem Paradebeispiel eines ITF, spielt HGV keine Rolle. Drei Hauptstädte, Zürich, Basel und Bern, gruppieren sich sternförmig um das kleine Olten herum, Fahrzeit jeweils im Bereich von 30 Minuten. Diese Gelegenheit wurde genutzt, um einen Kern-ITF zwischen diesen Städten mit 30-Minuten-Takt und integralem Umsteigen in Olten zu schaffen, auch wenn dies viele Ausbauten erforderte. Davon ausgehend wurden und werden Ausläufer nach Biel, Lausanne und Interlaken auf der Westseite sowie St. Gallen auf der Ostseite gestaltet. In dieser Gesamtheit stellt sich das Schweizer Netz wenig vermascht dar, sondern zieht sich eher linienförmig nördlich der Hochalpen entlang. In einer überwiegend linienförmigen Struktur ist ITF besonders einfach realisierbar.⁵

Auch das langgezogene Japan weist eher linienförmige, nur wenig vermaschte Netzstrukturen auf, doch steht hier, im Unterschied zur Schweiz, der HGV (Shinkansen) im Vordergrund. Die übrigen Fern- und Regionalverkehrsstrecken werden als Feederverkehre in den Takt der Shinkansen eingehängt. Hinzu kommen Nahverkehre in den Metropolregionen, die so eng getaktet sind, dass mit ihnen automatisch alles „integral“ ist (eine Implikation des sogenannten Mohring-Effektes).

Das HGV-Netz in Frankreich geht sternförmig und damit jeweils linienförmig von Paris aus in alle Richtungen.⁶ Zwischen diesen Linien gibt es kaum Querverbindungen. Der eigentlichen Herausforderung für einen ITF in Frankreich, dem Umsteigen im Superknoten Paris, stellt man sich nicht. Vielmehr stellt man sich auf den Standpunkt, dass in Paris auch diejenigen Passagiere, die vielleicht noch weiterfahren wollen, sich etwas

5 Wenn ein „Netz“ wirklich nur aus einer Linie besteht, dann ist ein einfacher Taktverkehr auf dieser Linie auch schon identisch mit einem „integralen Taktverkehr“.

6 Insb. Paris-Lyon-Marseille; sodann Paris-Brüssel oder Paris-Lille-London; Paris-Straßburg-Karlsruhe; Paris-Rennes; Paris-Bordeaux.

Zeit nehmen sollten – „Paris ist eine Messe wert“. Allerdings gibt es einige durchgebundene Züge (z.B. Marseille-Paris-Brüssel).

Wenn es überhaupt ein Land gibt, bei dem ein vermaschter ITF eine Herausforderung darstellt, der man gewillt ist, sich zu stellen, dann ist es Deutschland. Drei Nord-Süd-Achsen⁷ kreuzen sich mit vier Ost-West-Achsen⁸. Für Reisen zwischen den großen Metropolen (Berlin, Hamburg, München, Köln, Düsseldorf, Frankfurt) sowie einigen anderen Städten braucht man kaum umzusteigen, zumal DB Fern diese Hauptachsen in unterschiedlichen Kombinationen abfährt. Zwischen diesen Zügen gibt es nur wenige wirklich bedeutende Umsteigepunkte wie Mannheim, welcher aktuell gut funktioniert.

Um diese primären SPFV-Linien herum gibt es einen Kranz sekundärer SPFV- oder Regionallinien, die weitere Städte von außen anbinden und teilweise schon mit bestellten SPNV-Linien konkurrieren.⁹ Außerdem gibt es einige wenige wichtige Querverbindungen.¹⁰ Deutschland-Takt ist vor allem das Bestreben, die primären Linien mit diesen sekundären Linien und dem Nahverkehr i.e.S. zu einem ITF zu verknüpfen. Wenn dies gelänge, wäre das Verknüpfungspotenzial des ITF besonders hoch.

Die Fahrzeiten zwischen benachbarten Hauptknoten liegen zwar in ähnlichen Größenordnungen, variieren aber doch recht stark (deutlich stärker als in der Schweiz), was die erste Herausforderung eines ITF in Deutschland darstellt. Die zweite besteht darin, dass zwischen einigen Hauptknoten weitere wichtige Städte liegen, an denen in der Regel gehalten werden soll (genannt seien Stuttgart, Halle/Leipzig, Düsseldorf, Dortmund, Bremen).

Die dritte Herausforderung besteht darin, dass man aber auch HGV-Verbindungen zwischen weit entfernten Metropolen braucht, die eben nicht überall halten dürfen und daher aus dem ITF wenigstens teilweise herausgehoben werden müssen. Sie müssen in der Lage sein, andere ITF-Züge zu überholen. Im aktuellen Programm der DB Fern werden diese

7 Hamburg-Düsseldorf/Köln-Frankfurt-Mannheim-Basel; Hamburg-Hannover-Fulda-Nürnberg-München; Hamburg-Berlin-Erfurt-Nürnberg-München.

8 Berlin-Hannover-Düsseldorf/Köln; Erfurt-Fulda-Frankfurt; München-Mannheim-Frankfurt; München-Nürnberg-Frankfurt.

9 So ist im Nordosten des Landes die Linie Schwerin-Rostock-Stralsund in einem Dreieck mit Hamburg und Berlin einzubinden. Ferner die Städte Kiel, Lübeck/Kopenhagen, Dresden, Kaiserslautern/Saarbrücken, Aachen/Brüssel. Doch spielt das Umsteigen in diesen Städten eine untergeordnete Rolle.

10 So Stuttgart-Nürnberg und Hannover-Magdeburg-Halle/Leipzig.

Züge als „Sprinter“ bezeichnet, zum Beispiel Frankfurt-Berlin ohne Zwischenhalt.

Angesichts dieser Herausforderungen müssen im Deutschland-Takt wesentliche Abweichungen von einem idealen ITF hingenommen werden. Im Zentrum steht, wie früher schon, das Spannungsfeld zwischen ITF und echtem HGV (ab 300 km/h). Große Erfolge feiert der SPFV immer dann, wenn neue HGV-Strecken in Betrieb genommen und die Fahrzeiten zwischen weit entfernten Metropolen stark vermindert werden können. So führte die Eröffnung der neuen HGV-Strecke München-Berlin im Jahr 2017 zu einer Verdopplung des Verkehrsaufkommens auf 4,4 Millionen Reisende bei gleichzeitiger Abnahme der Flugzahlen um 1,2 Millionen Reisende auf dieser Verbindung (Redaktion Eurailpress, 2018). Auch im Ausland gab es die großen Erfolge nach der Eröffnung neuer HGV-Strecken über weite Distanzen (Paris-Lyon-Marseille; Mailand-Rom (Cascetta & Coppola, 2015); Japan (Takatsu, 2007)).

Berichte über ähnliche Erfolge nach der Einrichtung von Umsteigemöglichkeiten gibt es nicht. Hingegen ist bekannt, dass das Umsteigen von Fahrgästen generell als besonders lästig empfunden wird (Brenck et al., 2008). In Karlsruhe wird sogar bei der Gestaltung des innerstädtischen Straßenbahnnetzes darauf geachtet und viel Geld dafür investiert, dass möglichst viele Verbindungen ohne Umsteigen möglich sind. Karlsruhe kann darüberhinaus umsteigefreie Fernverkehrsverbindungen sowohl nach Berlin als auch nach Paris vorweisen. Nach Paris gibt es dabei genau einen Zwischenhalt (Straßburg) und die Fahrt dauert zweieinhalb Stunden. Nach Berlin sind es zehn Zwischenhalte und die Fahrt dauert fünfeinhalb Stunden. Somit dauert die Fahrt nach Berlin 120 % länger als die nach Paris, dabei ist die Entfernung nur gerade mal um 20 % Luftlinie und 40 % Eisenbahnlinie länger.

Es stellt sich daher die Frage, ob die Bedeutung von Umsteigemöglichkeiten derzeit nicht etwas überschätzt und die Bedeutung schneller, weiter Direktverbindungen mit echtem HGV unterschätzt werden. Die beachtlichen konkreten Planungen für einen Deutschland-Takt-Zielfahrplan (SMA und Partner et al., 2021) stellen, wie der aktuelle Fahrplan der DB Fern, einen Kompromiss zwischen den idealen des ITF und des HGV dar. Durch Einbeziehung der für den Zielfahrplan benötigten Infrastruktur-Maßnahmen kann dieser Fahrplan etliche Fahrzeitverkürzungen im Vergleich zum aktuellen Fahrplan der DB Fern erreichen. Die größten Fahrzeitverkürzungen gibt es wieder einmal durch die geplante Beschleunigung von Strecken (HGV Erfurt-Fulda, Ertüchtigung Dresden-Berlin). Aber auf drei Stunden Fahrzeit von Karlsruhe nach Berlin, vergleichbar mit dem französischen HGV, wird man auch damit nicht kommen.

Die vierte Herausforderung des Deutschland-Takts besteht darin, dass noch mehr Städte an den Fernverkehr angebunden werden sollen, sodass sowohl das Takterfordernis im Nah- und Regionalbereich noch komplexer als auch das Spannungsverhältnis zum HGV weiter verstärkt wird. Die fünfte Herausforderung besteht darin, dass die Länder und ihre Aufgabenträger immer mehr Regionalisierungsmittel vom Bund fordern und als Bestandteil der Klimapolitik auch durchsetzen, mit denen immer mehr Nah- und Regionalzüge bestellt werden, die die Hauptknoten zusätzlich belasten.

3. Wettbewerb im und zum Deutschland-Takt

Bei der Umsetzung eines Deutschland-Takt-Fahrplans wird sich die Frage stellen, ob und wie dieser mit intramodalem Wettbewerb im SPfV verträglich ist oder wettbewerbsfähig gestaltet werden kann. Dabei ist zu berücksichtigen, dass der Marktanteil der DB Fern immer noch bei 99 Prozent liegt. Allerdings gibt es seit der Markttöffnung im Jahr 1994 auch einzelne Züge, die von Wettbewerbern betrieben werden.

Im Kontext eines Deutschland-Takts ist es grundsätzlich möglich, dass Wettbewerber sich als Teil des Takts, der im Übrigen von der DB Fern betrieben wird, verstehen und von den Kunden auch so wahrgenommen werden. Es ist aber auch möglich, dass sich Wettbewerber explizit vom Deutschland-Takt abgrenzen, um als Alternative zum Takt wahrgenommen zu werden. Im ersten Fall kann man von „Wettbewerb im Deutschland-Takt“, im zweiten von „Wettbewerb zum Deutschland-Takt“ sprechen.

3.1 Wettbewerb im Deutschland-Takt

Im Folgenden werden drei Varianten des Wettbewerbs im Deutschland-Takt unterschieden, mit aufsteigendem Grad staatlicher Organisation.¹¹ In Variante 1 wird noch auf eine Planungsbehörde verzichtet. Variante

11 Solche drei Varianten wurden in der KCW-Studie von Berschin et al. (2019) entwickelt. Siehe auch Monopolkommission (2017, Kapitel 2.7.5, 2019, Kapitel 3.5, 2021, Kapitel 5.3) und die etwa gleichzeitig zu diesem Artikel entstandene Studie von Böttger (2021). Ideen aus diesen Quellen werden im Folgenden großzügig verwendet und natürlich wieder variiert.

3 entspricht hingegen weitgehend dem durchorganisierten Modell des SPNV. Dazwischen steht Variante 2, in der im Gegensatz zu Variante 3 noch keine staatlichen Gelder für die Bestellung von Verkehren eingesetzt werden.

Variante 1: Informeller Deutschland-Takt

In einem Modell mit sehr geringer formeller staatlicher Organisation würde die DB Fern einfach die Trassen nachfragen, die einen Deutschland-Takt ergeben, und in der Regel auch bekommen. Wettbewerber können jedoch selektiv eindringen, indem sie einzelne Trassen oder Pakete von Trassen nachfragen, die in das Deutschland-Takt-Gefüge passen. Zum Beispiel könnten zwei Unternehmen jeweils um eine Stunde versetzt im Zweistundentakt fahren.¹² Aus Sicht der DB Fern würden sich die Wettbewerber wie „Parasiten“ unerwünscht in ihre eigenen Netzzusammenhänge einklinken. Aus Sicht der Wettbewerber sind dies dann aber Netzzusammenhänge des Deutschland-Takts, nicht der DB, und die Möglichkeit sich darin einzuklinken neutralisiert teilweise die bisherige Markteintrittsbarriere, welche sich aus der hohen Netzbildungsfähigkeit des dominanten Anbieters ergibt.¹³

Natürlich wäre es auch denkbar, dass die DB Fern bestimmte, weniger attraktive Linien gerne an Wettbewerber abgibt und sogar entsprechende Vereinbarungen mit ihnen trifft, wie es auch aus dem Güterverkehr bekannt ist. Hier sollen aber nicht solche Kooperationsmodelle, sondern echte, konkurrierende Wettbewerber betrachtet werden.

Eintritt von „nicht-freundlichen“ Wettbewerbern in den Deutschland-Takt wäre auch ohne zwingende Durchtarifizierung möglich. Kunden würden dann mehrere Fahrkarten kaufen und könnten in den Knoten zwischen den Zügen unterschiedlicher Anbieter gut wechseln. Die Idee des Deutschland-Takts geht zwar bisher mit der Idee einer Durchtarifizierung einher, gleichzeitig entwickeln sich aber auch weitergehende Konzepte: Sogar für die Kombination von Fahrkarten verschiedener intermodaler Anbieter (zum Beispiel für Flüge und Bus- oder Taxifahrten) werden Anwendungssoftware (Apps) und Plattformen anvisiert, die Kombi-Fahrkar-

12 Solche Modelle gibt es derzeit zum Beispiel auf der Strecke Köln-Brüssel zwischen DB Fern und Thalys und auf der Strecke Hamburg-Bremen zwischen DB Fern und Metronom.

13 Zur Industrieökonomie der Netzbildungsfähigkeit siehe die theoretische Analyse von Hendricks et al. (1997).

ten und bei stärkeren Kooperationswünschen der Anbieter auch gemeinsame Tarife ermöglichen (z.B. Trainline.com Limited, 2020). Solche Technologien könnten auch innerhalb des Deutschland-Takts anbieterübergreifend genutzt werden. Ein Deutschland-Takt mit verschiedenen Anbietern wäre sogar das ideale Biotop für die Entstehung solcher unabhängigen Plattformen. Die dann sicherlich sofort aufkommende Warnung (von interessierter Seite), dass internationale Großkonzerne das attraktive Geschäft der Kundenvermittlung an sich ziehen würden, sollte man getrost ignorieren.¹⁴

Allerdings sind einige regulatorische Eingriffe der Bundesnetzagentur zur Gestaltung von Bahntarifen im Kontext des Deutschland-Takts denkbar und sicherlich wünschenswert. Für Bahnkunden wäre es besonders interessant, wenn beim Wechsel des Anbieters das Ticket des einen Unternehmens als Teilzahlung für dasjenige des anderen Unternehmens (welches vielleicht eine Stunde später fährt) akzeptiert wird, wobei die Verrechnung zwischen den Unternehmen dann im Hintergrund geschieht.¹⁵

Sollte es im Rahmen des Deutschland-Takts politisch gewünscht sein, einheitliche Tarifstrukturen und Fahrpreise für alle Anbieter im Taktsystem vorzugeben, müsste eine Behörde gegründet werden, die dies organisiert. Das Neue daran wäre weniger die Vereinheitlichung der Fahrpreise verschiedener Anbieter als die Tatsache, dass dann der dominante Anbieter DB Fern in seiner Endkundenpreisbildung reguliert würde. Das ist bisher nicht der Fall (im Gegensatz zu Frankreich), und es gibt derzeit auch keine politischen Bestrebungen in diese Richtung.

Variante 2: Konstruktion und Vermarktung der Deutschland-Takt-Trassen durch eine Behörde

Für die Entwicklung des Deutschland-Takts sind auch Varianten denkbar, die im Zuge eisenbahnrechtlicher Anpassungen stärker von einer Behörde strukturiert werden (vgl. Monopolkommission, 2017, S. 235ff., 2019, S. 213ff., 2021, S. 112f., sowie Böttger, 2021, Kapitel 4). So könnte eine Deutschland-Takt-Planbehörde den Takt konzipieren: Welche Knotenbegegnungen sind vorgesehen und welche Takttrassen zwischen den Knoten ergeben sich daraus? Diese Trassen müssten zwischen DB Netz und Plan-

14 So ist es keineswegs der Fall, dass Flugsuchportale einen wesentlichen Teil der Wertschöpfung des Flugsektors an sich ziehen.

15 Die Vorteile der Transferierbarkeit von Tickets zeigen Brand und Sieg (2020).

behörde verhandelt und vorkonstruiert werden. Die Planbehörde könnte die Trassen kaufen und sie dann unter von ihr bestimmten Bedingungen an Verkehrsunternehmen vermarkten. Diese Bedingungen können sich auf Merkmale der Servicequalität und der Tarifstrukturen erstrecken (und auch Obergrenzen für die Fahrpreise enthalten, wenn dies gewünscht wäre).

Die Vermarktung würde die Form eines Ausschreibungswettbewerbs für die Deutschland-Takt-Trassen(-Pakete) annehmen, wobei die normalen Trassenpreise im ersten Schritt das Mindestgebot darstellen würden. Im zweiten Schritt könnte auf diese Mindestpreise auch verzichtet werden, was schon eine gewisse Subventionierung beinhalten würde (siehe Berschin et al., 2019, S. 24). Für die Variante 2 sei jedoch ausgeschlossen, dass öffentlicher Gelder zur Bestellung von Verkehren eingesetzt werden.

In einem solchen Rahmen wäre eine gezielte Förderung des Wettbewerbs im Deutschland-Takt möglich. So könnten nicht marktbeherrschende Unternehmen bei der Vergabe dieser Trassen bevorzugt werden oder ein bestimmter Anteil der Trassen für sie reserviert werden.

Der Sinn eines organisierten Wettbewerbs würde darin bestehen, dass Kunden und Planbehörde Preis- und Qualitätsvergleiche zwischen den verschiedenen Unternehmen anstellen und darauf reagieren können. Betrachten wir wieder den Fall, dass eine Strecke von zwei Unternehmen alternierend bedient wird. Die Kunden können dann schnell reagieren, indem sie den besseren Anbieter bevorzugen. Die Planbehörde kann in langfristigen Intervallen reagieren, indem sie den besseren Anbieter in der Zukunft bei der Vergabe von Deutschland-Takt-Trassen bevorzugt, wofür eine geeignete, transparente Form der Bewertung und Honorierung zu finden wäre.¹⁶

In allen Varianten sind Wettbewerber darauf angewiesen, für sich selbst ein Mindestpaket zusammenpassender Trassen zu erhalten, welches eine wirtschaftlich sinnvolle Investitions-, Umlauf- und Personalplanung ermöglicht. Dieses Erfordernis wird auch im Rahmen des Deutschland-Takts weiterhin eine Markteintrittsbarriere darstellen. Im Kontext der informellen Variante 1 wäre ein von der Bundesnetzagentur koordinierter Verhandlungsprozess vorstellbar, um solche Trassenpakete in Hinblick auf die konkreten Geschäftsmodelle der Wettbewerber zu schnüren.

Im Kontext eines Ausschreibungswettbewerbs (Variante 2) wäre eine kombinatorische Auktion vorstellbar, bei der die Unternehmen auf für sie

16 Auf mögliche Koordinationsprobleme zwischen Wettbewerbern in einem ITF weisen Brand und Sieg (2022) hin.

wirtschaftlich sinnvolle Kombinationen von Trassen aus der Menge der vorkonstruierten Deutschland-Takt-Trassen bieten können. Es gibt aber in der Wirtschaft nur wenige Praxisbeispiele für erfolgreiche komplexe kombinatorische Auktionen. Schon die komplexe Anforderung an die Unternehmen, in einer kombinatorischen Auktion sinnvoll dynamisch zu bieten (um etwa das Risiko zu minimieren, Zuschläge für unbrauchbare Einzeltrassen zu erhalten) würde ihrerseits eine Markteintrittsbarriere darstellen (Tanner & Mitusch, 2011).

Alternativ könnte im Kontext der Variante 2 die Deutschland-Takt-Planbehörde selbst verschiedene Deutschland-Takt-Teilnetze zusammenstellen und als Pakete vermarkten. Dafür müsste sie Teilnetze schaffen, die nach ihrer Meinung (sicherlich nach Anhörung von Marktteilnehmern) eine wirtschaftlich sinnvolle Investitions-, Umlauf- und Personalplanung ermöglichen. Das würde die Markteintrittsbarrieren für Wettbewerber deutlich verringern.

Nachteil solcher behördlichen Planungsvorgaben ist allerdings, dass die entsprechenden Gestaltungsaufgaben nicht mehr von den Unternehmen im Wettbewerb gelöst werden. Dann ist davon auszugehen, dass sie langfristig weniger gut gelöst werden, denn der Staat ist kein guter Unternehmer und auch nicht frei von Interessen und Einflussnahmen, die einer effizienten Gestaltung entgegenarbeiten. Dies waren die zentralen Grundeinsichten, welche zur deutschen Bahnreform von 1994 und zum ersten Eisenbahnpaket der Europäischen Gemeinschaft von 2001 führten.

Allerdings sind auch Unternehmen, die dauerhaft vor Wettbewerb geschützt sind, selten gute Unternehmer, weder aus gesamt- noch aus einzelwirtschaftlicher Sicht. Daher wird die Aufgabe darin bestehen, einen guten Kompromiss zwischen Planungsvorgaben, die den Markteintritt von Wettbewerbern erleichtern, und der Ermöglichung tiefgehender unternehmerischer Gestaltung zu finden. Hier könnte das im Telekommunikationssektor entwickelte dynamische Konzept der „Ladder of Investment“ einen Gedankenanstoß geben, wonach der Staat zunächst stark intervenieren sollte, um Markteintritte mit recht flachen Geschäftsmodellen zu stimulieren, und später die Eingriffstiefe reduzieren sollte, in der Hoffnung, dass diese Unternehmen dann nicht wieder aussteigen, sondern ein vertieftes Geschäftsmodell entwickeln (Cave, 2006).¹⁷

17 Zur Ladder eher kritisch in Hinblick auf den Telekommunikationssektor Knieps und Zenhäusern (2007, S. 262).

Variante 3: Mit öffentlichen Geldern bestellte Fernverkehre

Es gibt weitergehende verkehrspolitische Vorstellungen zur Gestaltung eines Deutschland-Takts, die nur mithilfe öffentlicher Subventionen möglich sein werden: Engerer Takt, günstigere Fahrpreise, mehr Angebote auch an den Tagesrandzeiten und in der Nacht, mehr Städte anbinden, den Deutschland-Takt noch enger mit dem regionalen SPNV integral verknüpfen.

Wenn der Staat den Betrieb bestimmter Trassen bestellen und subventionieren möchte, ist Ausschreibungswettbewerb unumgänglich. Trassen und Bestellerentgelte gehen dann an denjenigen Bieter, der die geringste Subvention fordert. Dies erfordert eine Deutschland-Takt-Planbehörde, wie schon für Variante 2 skizziert, so dass sich das Marktmodell des Deutschland-Takts stark dem des Ausschreibungswettbewerbs im SPNV annähert. Allerdings vergeben die Aufgabenträger des SPNV oft (nicht immer) den gesamten Verkehr ihres Gebietes an einen einzigen Bieter. Die Deutschland-Takt-Planbehörde muss Teile des Takts wirtschaftlich sinnvoll zurechtschneiden und separat ausschreiben, damit Wettbewerb um bestellte Verkehre entstehen kann.

Da Steuergelder involviert sind, ist die Deutschland-Takt-Planbehörde viel stärker als bei Variante 2 darauf angewiesen, dass sich wirklich in einem überschaubaren Zeitraum signifikanter Wettbewerb entwickelt. Mit großangelegten Manövern, die auch zusätzliches Geld kosten werden, muss die Behörde Markteintritte Dritter fördern und so Wettbewerb um die bestellten Verkehre generieren.¹⁸ Mögliche Adressaten könnten die französische SNCF, die italienische Italo NTV, die deutsche Flixbus oder Unternehmen, die bisher im deutschen SPNV tätig sind, sein.

Für potenzielle Wettbewerber wäre die Aussicht auf öffentliche Bestellungen und ordentliche Ausschreibungswettbewerbe schon ein wichtiger Anreiz für Markteintritte. Aber das wird vermutlich nicht reichen. Wichtig ist, wie schon bei Variante 2 angesprochen, ein Zuschuss von Ausschreibungslosen, der den Markteintritt erleichtert, ggf. in Verbindung mit gezielten Bevorzugungen der Wettbewerber. Zudem könnte in Variante 3 die Deutschland-Takt-Planbehörde einen Fahrzeugpool für Züge des SPNV unterhalten, der seine Fahrzeuge für die Vertragsdauer an die Betreiber

18 Es gibt zwei andere Optionen: Die DB Fern könnte in verschiedene Unternehmen aufgespalten werden, die separat privatisiert werden, um sich danach Wettbewerb zu liefern. Oder man verzichtet auf Wettbewerb und reguliert die DB Fern unmittelbar (wie in Frankreich). Beide Wege halte ich für nicht empfehlenswert und möchte sie aus Platzgründen hier nicht diskutieren.

vermietet, wie dies in einigen Bundesländern im SPNV geschieht, um dort den Wettbewerb zu fördern (vgl. Monopolkommission, 2021, Kapitel 6.3, Böttger, 2021, Kapitel 3.3.5).¹⁹ Nachteil dieser Vorgehensweisen wäre allerdings, dass immer weniger Entscheidungen unternehmerisch im Wettbewerb getroffen werden. Fahrzeugpools sollten daher nur in überschaubarem Umfang geschaffen und ohne Verpflichtung zu ihrer Nutzung angeboten werden. Firmen wie SNCF und Italo NTV würden vermutlich eher mit ihrem eigenen Zugmaterial aus französischer Produktion eintreten. Für sie und alle anderen Wettbewerber sollte aber in Variante 3 die Deutschland-Takt-Planbehörde Unterstützung für die Logistik eigener Zugflotten anbieten (Werkstätten etc.). Eine weitere Form der Wettbewerbsförderung könnte darin bestehen, regulierten Zugang zu Sitzplatzkapazitäten beim marktbeherrschenden Unternehmen zu gewähren. Damit könnten Wettbewerber auch mit beschränkten Angeboten Verbundvorteile ohne hohe Markteintrittskosten realisieren. Dieses Zugangsrecht könnte an das Vorhandensein von eigenen Zügen mit passenden Umsteigemöglichkeiten gebunden werden.²⁰

3.2 Wettbewerb zum Deutschland-Takt

Wenn ein Deutschland-Takt umgesetzt ist, können daneben auch noch andere Angebote des SPNV existieren, die nicht in den Deutschland-Takt integriert sind. Denn der Zugang zum Schienennetz bleibt grundsätzlich frei.

Wenn sich der Deutschland-Takt nach der im vorigen Abschnitt skizzierten informellen Variante 1 entwickelt, verschwimmt die Grenze zwischen „Wettbewerb zum“ und „Wettbewerb im“ Deutschland-Takt. Sie könnte aber im Marketing der Unternehmen eine Rolle spielen, wenn diese ihr Angebot als „Alternative zum Deutschland-Takt“ anpreisen. Je mehr der Deutschland-Takt institutionell durch eine Planbehörde vorstrukturiert wird (Varianten 2 und 3), desto klarer werden andere Angebote von diesem unterscheidbar sein.

Man kann zwischen „direktem Wettbewerb“ und „Randwettbewerb“ zum Deutschland-Takt unterscheiden. Während direkter Wettbewerb dem

19 Ein Beispiel hierfür ist der niedersächsische SPNV-Aufgabenträger LNVG.

20 Die Idee des regulierten Zugangs zu Sitzplatzkapazitäten wurde jüngst von Eisenkopf und Knorr (2021) in die Diskussion eingebracht und wird hier in einer etwas restriktiveren Form aufgegriffen.

Deutschland-Takt mit sehr ähnlichen Angeboten entgegentritt, versucht Randwettbewerb eher, den Angeboten des Deutschland-Takts auszuweichen und Marktnischen zu finden.

Direkter Wettbewerb zum Deutschland-Takt

Direkter Wettbewerb zum Deutschland-Takt liegt dann vor, wenn ein Wettbewerber weitgehend parallel zu einem Zug des Deutschland-Takts mit ähnlichem Zugmaterial operiert. Durch Preis- und Qualitätsgestaltung oder einfach durch die von ihm angebotene zusätzliche Beförderungskapazität versucht er, dem Deutschland-Takt Kunden streitig zu machen. Zugleich kann er versuchen, an ausgewählten Knoten an den Deutschland-Takt anzudocken. Direkter Wettbewerb würde das Bahnangebot insgesamt erweitern und den Deutschland-Takt unter einen Preis- und Qualitätsdruck setzen; er könnte dadurch zu einer Verbesserung der Deutschland-Takt-Angebote beitragen.

Dies könnte sogar zu der Befürchtung führen, dass direkter Wettbewerb auf ausgewählten Strecken ein eigenwirtschaftlich aufgesetztes Taktsystem so stark unter Druck setzen würde, dass es in seiner Existenz gefährdet wäre. Als nötig empfundene Quersubventionierungen zwischen rentablen und unrentablen Verkehren wären dann nicht mehr leicht aufrechtzuerhalten (vgl. Knieps, 2007, S. 143ff.). Eine mögliche Konsequenz dieser Überlegung könnte also sein, dass direkter Wettbewerb zu verbieten sei.

Allerdings greift diese einfache Überlegung zu kurz. Denn ein direkter Wettbewerber steht auch selbst unter starkem Konkurrenzdruck durch die Deutschland-Takt-Angebote. Insbesondere aufgrund der Netzbildungsfähigkeit der Deutschland-Takt-Angebote kann diese Resistenz so stark sein, dass ein direkter Wettbewerber am Markt nicht bestehen kann (vgl. Hendricks et al., 1997). Das gilt erst recht, wenn die Deutschland-Takt-Angebote auch noch subventioniert wären (Variante 3).²¹ Zudem werden die Infrastrukturkapazitäten insbesondere im Bereich der Knoten für einen perfekt eingetakteten direkten Wettbewerb sehr knapp sein. Denn profitable direkte Wettbewerbsangebote sind nur auf den stark nachgefragten

21 So hat es bis heute keinen Markteintritt in Konkurrenz zum subventionierten SPNV gegeben, obwohl dies möglich wäre. Auf einigen Strecken wäre bei Abwesenheit von Subventionen Wettbewerb sicherlich möglich (zum Beispiel Hamburg-Lübeck).

Hauptstrecken zwischen den Großstädten denkbar, und gerade dort sind die Infrastrukturkapazitäten knapp.

In allen Branchen versuchen neue Marktteilnehmer, sich von den bisherigen Angeboten zu unterscheiden, um bestimmte Kundensegmente besser bedienen zu können und den Konkurrenzdruck abzumildern. Aus diesem und den anderen genannten Gründen ist "perfekter" direkter Wettbewerb zum Deutschland-Takt unwahrscheinlich. Aber dort, wo der Deutschland-Takt ein Stundentakt ist, könnte ein Wettbewerber jeweils in die halbe Stunde dazwischen gehen. Dann sind die Personenbahnhöfe und die dazu gehörigen Knoteninfrastrukturen frei, und die Passagiere können unter Inkaufnahme einer halbstündigen Wartezeit immer noch recht gut in die Züge des Deutschland-Takts umsteigen oder aus diesen kommen.

Die mit dem Deutschland-Takt verbundenen Infrastrukturausbauten und Vereinheitlichungen der Betriebsprogramme verbessern eigentlich die Voraussetzungen für einen zeitlich versetzten, „imperfekten direkten“ Wettbewerb zum Deutschland-Takt. Denn es sollen ja möglichst viele Züge (auch des SPNV) auf die Taktstunde konzentriert werden, und die Knotenkapazitäten sollen für diese hohe Belastung ausgebaut werden. Die Folge wird sein, dass in der Zwischenzeit in den Knoten viel freie Infrastruktur zur Verfügung stehen wird. Da aktuell die meisten Infrastrukturengpässe in den Knoten zu finden sind, sollte ein Wettbewerber zwischen den Taktwellen auch die benötigten Trassen finden können. Allerdings werden zwischen den Taktwellen die Kanten und die von den Bahnhöfen etwas weiter entfernten Bereiche der Knoten auch vom Güterverkehr stark genutzt, so dass die Wettbewerber hier auf Infrastrukturengpässe treffen können. Insbesondere dort, wo der Deutschland-Takt zum Halbstundentakt verdichtet wird, wird es für einen Wettbewerber sehr schwierig, sich noch dazwischen zu setzen, zumal das Deutschland-Takt-Angebot dann schon sehr erdrückend in Hinblick auf die vorhandene Nachfrage sein wird.

Ein Wettbewerber könnte versuchen, einen Geschwindigkeitsvorteil gegenüber dem Deutschland-Takt dadurch zu erlangen, dass er nicht an allen Knoten hält und dass seine Halte von kürzerer Dauer sind. Das kann aber auf der langen Strecke nur funktionieren, wenn er die Takt-Züge überholen kann. Zudem müsste an den Knoten der zeitliche Versatz zum Takt variabel sein, nicht immer genau zur halben Zeit. Ob die Knoten- und Streckenkapazitäten des Netzes dies hergeben, wird vom Einzelfall abhängen. Als flächendeckendes Geschäftsmodell ist es schwer vorstellbar.

Randwettbewerb zum Deutschland-Takt

Wettbewerber werden daher vorwiegend dort eintreten, wo der Deutschland-Takt Defizite und Lücken aufweist. Dabei werden sie zeitlich oder in der Streckenführung oder in den von ihnen angesprochenen Personengruppen den Deutschland-Takt-Angeboten tendenziell ausweichen. Dies ist der Randwettbewerb.

Die aktuellen Konkurrenzangebote von FlixTrain zur DB Fern sind Randwettbewerb. Einige Linienführungen von FlixTrain weichen von denen der DB Fern in bestimmten, aber wichtigen Abschnitten ab. Ein Beispiel hierfür ist die Strecke Stuttgart – Berlin (FLX 10) über Hannover, Göttingen und Frankfurt (die vergleichbare ICE-Linie der DB Fern hält nicht in Hannover). Andere Linien laufen zwar parallel (zum Beispiel Hamburg – Köln), doch adressieren sie nach Preis und Gestaltung ein anderes Publikum. Die Züge von FlixTrain fahren viel seltener als die Züge der DB Fern (je nach Strecke ein bis drei Zugpaare täglich) und stellen schon aus diesem Grund ein Randphänomen dar. Dennoch kann auch diese beschränkte Form des Wettbewerbs für die Kunden, den Markt und die zukünftige Marktentwicklung bedeutsam sein und auch einen heilsamen Preis- und Qualitätsdruck auf die Deutschland-Takt-Angebote ausüben.

4. Eigenwirtschaftlichkeit versus Besteller-Prinzip im Deutschland-Takt

„Economies of density” (Dichtevorteile) sind ein wesentliches Merkmal aller Netzwerksektoren: Die Fixkosten von Infrastrukturen, öffentlichen Verkehrsangeboten und anderen Diensten fallen umso mehr ins Gewicht, je dünner die Nachfrage ist. Bei verdichteter Nachfrage verteilen sie sich hingegen auf viele Köpfe, so dass die Eigenschaft des „natürlichen Monopols” verblasst. So herrscht im Internetsektor starker Infrastrukturwettbewerb im überregionalen „Backbone” vor, während in den lokalen Anschlussnetzen Regulierungsbedarf festgestellt wird. Der Postsektor erlaubt Gewinne und Wettbewerb in verdichteten Räumen, während in ländlichen Räumen selbst eine Grundversorgung („Daseinsvorsorge”) unwirtschaftlich ist. Die Standardlösung für einen solchen Fall sind organisierte, wettbewerblich gestaltete „Universaldienstverpflichtungen”: Alle Unternehmen des Sektors werden mit einer Abgabe belegt, deren Erlös dazu dient, die Bedienung der ländlichen Räume zu subventionieren. Die

subventionierten Dienste werden in Ausschreibungswettbewerben an die günstigsten Betreiber vergeben (Knieps, 2007, Kapitel 7).²²

Diese ökonomischen Grundprinzipien gelten auch für den öffentlichen Verkehr. Auf Verbindungen zwischen den Metropolen, wo sich Nachfrage konzentriert („Backbone“), sowie innerhalb der Metropolen selbst, also im Nahverkehr in und zwischen dicht besiedelten Stadtgebieten, sollten profitable Bahn- und Busangebote im Wettbewerb möglich sein. Grundsätzlich gilt daher das Prinzip der Eigenwirtschaftlichkeit, d.h. Subventionsfreiheit, auch im öffentlichen Verkehr.²³ Allerdings ist die Versorgung ländlicher Räume und dünn besiedelter Stadtgebiete schwierig. Nach dem Prinzip der Universaldienstverpflichtung sollten daher die Metropolverkehre mit einer Abgabe belegt werden, um die unprofitablen Nah- und Regionalverkehre zu subventionieren. Letztere sollten in öffentlichen Ausschreibungen vergeben werden.

Bei der Eisenbahn wurden diese Grundprinzipien teilweise umgesetzt. Der SPFV soll eigenwirtschaftlich zurechtkommen, während der SPNV subventioniert und (inzwischen überwiegend) wettbewerblich ausgeschrieben wird. Das Prinzip der Universaldienstverpflichtung wurde aber insofern abgeschwächt, als der SPFV nicht mit einer Abgabe zur Unterstützung des SPNV belastet wird. Vielmehr wird der SPNV direkt aus Steuermitteln subventioniert, durch die Regionalisierungsmittel des Bundes in Höhe von (vor Corona) 8 Mrd. Euro pro Jahr. Von diesen Geldern profitiert auch der SPFV, da der SPNV Kunden für den Fernverkehr einsammelt und verteilt, ebenso wie übrigens die städtischen Bus- und Bahnverkehre, die ebenfalls aus Steuermitteln subventioniert werden. Die Schieneninfrastruktur wird ohnehin weitgehend vom Staat finanziert, wobei der SPFV in besonderer Weise profitiert, da HGV-Strecken besonders teuer sind.²⁴ Daher kommt der Bund seiner Allgemeinwohlverpflichtung bei den Verkehrsangeboten des SPFV seit langem nach (vgl. Art. 87e Abs. 4 GG), und die Eigenwirtschaftlichkeit des SPFV ist nicht nur ökonomisch gut be-

22 Beispielspielhaft Postgesetz, Abschnitt 3. Ähnlich Telekommunikationsgesetz, Teil 9, doch ist dort eine Ausschreibung nicht vorgesehen, sondern nur eine Prüfung der Kosten des von der Bundesnetzagentur verpflichteten Anbieters.

23 Vgl. Personenbeförderungsgesetz, § 8.

24 Die Trassenpreise finanzieren nur etwa 40 % der Ausgaben der DB Netz AG, der Rest kommt direkt vom Bund. Zwei Drittel der Trasseneinnahmen kommen aus dem SPNV (BNetzA, 2021, S. 22) und somit zum großen Teil indirekt vom Bund, über die Regionalisierungsmittel. Dies entspricht auch in etwa dem Anteil des SPNV an den gesamten Trassenkilometern, die auf dem Netz gefahren werden (ebd., S. 16), doch nutzt der SPNV nicht die HGV-Fähigkeit der Infrastruktur.

gründet, sondern auch durch öffentliche Gelder satt finanziell unterfüttert.²⁵

Dennoch ist es für die DB Fern nicht leicht, ihre Eigenwirtschaftlichkeit betriebswirtschaftlich darzustellen. Sie erwirtschaftete zwar in den Jahren 2017 bis 2019 respektable Gewinne, doch ist aufgrund notwendiger Investitionsvorhaben unklar, ob dies so bleiben wird,²⁶ und das Problem der Unpünktlichkeit und der Zugausfälle bekommt sie nicht in den Griff. Sicherlich birgt diese Sparte des Staatskonzerns viele Effizienzreserven, doch werden sie nicht leicht zu heben sein. Möglicherweise werden im Rahmen des Deutschland-Takts zusätzliche Wettbewerbspotenziale realisiert, die auch die DB Fern besser auf Trab bringen könnten. Möglicherweise wird mit der Umsetzung des Deutschland-Takts auch zusätzliche Nachfrage generiert, die die Rentabilität des Gesamtsystems steigert. Doch sollte man sich nicht allzu große Hoffnungen machen: Substanzielle Angebotsausweitungen, Qualitätsverbesserungen oder Preissenkungen des SPFV wird es im eigenwirtschaftlichen Rahmen nicht geben.

Genau dies erwarten jedoch einige politische Akteure vom Deutschland-Takt und fordern dafür auch zusätzliche Steuergelder (Berschlin et al., 2019; Bundestagsfraktion Bündnis 90/Die Grünen, 2019). Bestimmt gibt es viele einzelne, spezifische Fälle, in denen sich staatliche Subventionen für den SPFV gut rechtfertigen lassen. Aber sobald es mehr als zwei werden, wird die Forderung nach einer systematischen Herangehensweise laut werden.

Sobald aber die Bereitschaft erkennbar wird, Geld dauerhaft und in größerem Umfang für SPFV-Bestellungen zur Verfügung zu stellen, etwa vergleichbar mit den Regionalisierungsmitteln des SPNV, wird eine Hydra organisierter Interessen auf den Plan treten, die auf einen ständigen Aufwuchs dieser Mittel drängen und, so ist zu befürchten, einen wachsenden Anteil dieser Mittel ineffizienten Verwendungen zuführen wird.

-
- 25 Erwähnenswert ist in diesem Kontext auch die Versorgung von Fernverkehrsstrecken mit schnellem Internet. Diese wird im Rahmen der Frequenzauktionen den Telekomunternehmen auferlegt, obwohl es sich bei vielen Streckenabschnitten nicht nur um „weiße Flecken“, sondern um völlig unbewohnte Gebiete handelt, durch die pro Stunde etwa 400 Personen im Zug befördert werden. Das ist richtig so, weil diese Unternehmen das besser können. Die Kosten in Form verringerter Auktionserlöse des Bundes sind eine weitere Subvention des SPFV.
- 26 Zur wirtschaftlichen Situation der DB Fern siehe Böttger und Mitusch (2016), den Bericht „Zur wirtschaftlichen Lage der DB Fernverkehr AG“ in FIS (2021) und Böttger (2021, S. 6).

Sobald der Druck der Eigenwirtschaftlichkeit wegfällt, wird ein erster Teil der öffentlichen Gelder direkt bei der DB Fern versickern – ohne signifikante Verbesserungen. Den nächsten beträchtlichen Teil wird sie brauchen, um bei nominell gleichbleibendem Angebot und faktisch gleichbleibenden internen Prozessen die Unpünktlichkeiten und Zugausfälle in den Griff zu kriegen. (Wenn dies gelingt, wird man es mit gewissem Recht als großen Erfolg feiern – die Frage ist nur, wie teuer es dann schon wird.) Jede weitergehende Forderung aus dem politischen Raum (Taktfrequenz, Städteanbindung, Fahrpreissenkung) wird sie sich fürstlich entlohnen lassen.

Ein Regulativ gegen solche Bestrebungen des marktbeherrschenden Unternehmens wird es nur dann geben, wenn es gelingt, signifikanten Wettbewerb im Deutschland-Takt zu etablieren. Aber ob das gelingt und wie stark die Wettbewerbsintensität sein wird, ist offen. Ebenso die Frage, welchen Aufwand die Deutschland-Takt-Planungsbehörde treiben muss, um den im Fall öffentlicher Subventionen unverzichtbaren Wettbewerb in das System hineinzudrücken. Es gibt nicht allzu viele potenzielle Wettbewerber, und deren Markteintrittskosten müssen erstmal eingespielt werden. Wenn sie an die Arbeitnehmer-Tarifverträge der DB gebunden sein werden, wird dies die Wettbewerbsintensität reduzieren. Der Nutzen des Wettbewerbs wird daher (neben den Anreizen zur Verbesserung der Servicequalität und zur Senkung der Fahrpreise) eher darin bestehen, die Subventionen zur Erfüllung eines gegebenen Bedarfs nicht völlig explodieren zu lassen und ihr langfristiges Wachstum etwas einzudämmen.

Hinzu kommen aber steigende Bedarfsanmeldungen und Ansprüche an den SPFV. Haupttreiber hierfür werden die Länder und ihre Kommunen und Aufgabenträger sein, die auf zusätzliche Städteanbindungen, höhere Taktfrequenzen und niedrigere Fahrpreise drängen werden. Auch könnten sie versuchen, Regionalverbindungen, die bisher vom SPNV betrieben werden, dem bestellten SPFV zuzuordnen, möglichst ohne eine entsprechende Reduktion ihrer eigenen Regionalisierungsmittel.²⁷ Diese wollen sie möglichst noch steigern, um die Nahverkehre auszubauen, mit der Folge, dass die Ansprüche an die Fernverkehrsverbindungen im Deutschland-Takt weiter steigen, ebenso wie die Belastung der Knoten.

27 Als die DB Fern in den 1990er Jahren die unprofitablen InterRegio-Züge einstellte, übernahmen die Aufgabenträger einen Teil dieser Fernverkehrsverbindungen zwischen kleineren Zentren im Rahmen des SPNV. Erfreut waren sie über diese Verschiebung der Kosten nicht und würden sie sicherlich gerne wieder rückgängig machen. Laut Böttger (2021, S. 1) gab es dafür 2002 eine finanzielle Kompensation vom Bund.

Eine Illustration für diese erwartbaren Vorgänge bieten die Erfahrungen mit dem subventionierten SPNV. Da der Bund für die Bahnreform von 1994 (Grundgesetzänderung) auf die Zustimmung der Länder angewiesen war, erstritten die Länder hohe Zuwendungen des Bundes für den SPNV, die seitdem im zustimmungspflichtigen Regionalisierungsgesetz vereinbart werden und stetig steigen, über alle Wirtschaftskrisen hinweg. Entsprechend nehmen auch die bestellten Leistungen zu, wie zum Beispiel anhand der steigenden Zugkilometer des SPNV sichtbar wird. In der Konsequenz hat sich der Nahverkehr, der vor der Bahnreform stark defizitär war, zu einer Cash-Cow des DB Konzerns gewandelt. Zwar haben sich im SPNV auch Ausschreibungswettbewerb und unternehmerische Wettbewerber etabliert und die DB Regio hat Marktanteile verloren, doch behauptet sie ihre Position und ihre Profitabilität in diesem subventionierten Geschäft. Der Wettbewerb erfüllt also eine gewisse Funktion, indem die Subventionen nicht völlig von einem Monopolisten abgeschöpft werden können,²⁸ doch konnte er nicht verhindern, dass das an sich schwierige Nahverkehrsgeschäft zu einer ausgezeichneten Geldquelle der DB wurde.

In der Gesamtheit ist die Entwicklung des SPNV dennoch positiv zu beurteilen, der heute ein attraktives und zeitgemäßes Angebot bietet, während er vor der Bahnreform von sehr schlechter Qualität war. Im Fernverkehr aber ist die heutige Ausgangssituation eine deutlich bessere, und dies hat strukturelle ökonomische Gründe, wie dargestellt wurde. Es spricht einiges dafür, das subventionstreibende Regime des SPNV nicht auf den SPNV zu übertragen, sondern die Trennung der Regime beizubehalten, um im Fernverkehrssegment einen gewissen Effizienzdruck zu wahren. Das würde aber bedeuten, auf einige Wünsche zu verzichten.

5. Fünf Vorschläge zur weiteren Entwicklung des Deutschland-Takts

Auf Grundlage der bisherigen Überlegungen folgen fünf Vorschläge, wie mit dem Deutschland-Takt weiter verfahren werden kann.

28 Angesichts der hohen Regionalisierungsmittel im Vergleich zu den potenziellen Fahrgeldeinnahmen des SPNV wäre die ökonomische Anreizsituation tatsächlich derart, dass ein unbestreitbarer Monopolist beliebig hohe Subventionen pro Zugkilometer verlangen würde, um die variablen und sprungfixen Kosten zu minimieren.

Vorschlag 1: Mit der Umsetzung des Taktfahrplans abwarten; keine bestellten Verkehre

Aktuell ist der Deutschland-Takt im Wesentlichen ein Investitionsprogramm des Bundes. Um die Engpässe im Netzkontext zu identifizieren und ihre Beseitigung sinnvoll aufeinander und auf die erwartete zukünftige Nachfrage abzustimmen, ist das zum Deutschland-Takt gehörige Konzept der „fahrplanbasierten Entwicklung der Schieneninfrastruktur“ gut geeignet. Bei dem Investitionsprogramm handelt es sich also um eine „no regret“- oder jedenfalls „low regret“-Strategie.

Wesentlich riskanter wäre hingegen die rigorose Umsetzung eines Deutschland-Takt-Fahrplans unter Einsatz umfangreicher Steuergelder zur Bestellung von Verkehren. Eine sinnvolle Planung, die sich an volkswirtschaftlicher Effizienz orientiert, könnte in einem solchen System durch die inhärenten anreiz- und politökonomischen Probleme leicht unterminiert werden. Das Problem liegt dabei weniger in der institutionellen Gestaltung einer hierfür benötigten Deutschland-Takt-Planbehörde als in dem Umfeld, in dem diese vermutlich agieren würde: Geringe Wettbewerbsintensität unter den potenziellen Verkehrsunternehmen; dominanter Anbieter DB Fern, der sich seiner Staatsnähe und vertikalen Integration mit den Infrastrukturbetreibern DB Netz und Station&Service wohl bewusst ist; mächtige und gut organisierte Länderinteressen, denen die Einbindung des Nah- und Regionalverkehrs und weiterer Städte in den Fernverkehr womöglich wichtiger ist als die Qualität des Fernverkehrs über weite Distanzen.

Daher ist es wohl ratsam, die bisherige, vorsichtig tastende Strategie zur Realisierung des Deutschland-Takts beizubehalten: Der Bund überlegt sich, wie zwischen den Idealen des ITF und des HGV ein guter SPFV in Deutschland einmal aussehen könnte (der auch adäquate Netzkapazitäten für den Schienengüterverkehr vorsieht (Wissenschaftlicher Beirat beim BMVI, 2021)). Studien dafür liegen vor. Darauf basierend wird der fahrplanbasierte Infrastrukturausbau in Angriff genommen. Gleichzeitig lässt man aber offen, wie und wieweit ein angedachter Taktfahrplan tatsächlich umgesetzt wird. Man schafft lediglich die Voraussetzungen. Gleichzeitig kann man informell, mit wenigen rechtlich-institutionellen Änderungen, auf eine maßvolle Realisierung entsprechender Fahrpläne hinarbeiten.

Zudem können Überlegungen zu einem „Deutschland-Tarif“ weiterverfolgt und umgesetzt werden. Einige Regulierungen zur gegenseitigen Anerkennung oder Anrechnung von Fahrscheinen können sinnvoll sein. Die Ausstellung durchgebundener oder kombinierter Fahrscheine kann man

privaten Portalen überlassen, die auch intermodal verknüpfen können; man sollte aber die Voraussetzungen dafür schaffen.

Dies entspricht der Vorgehensweise der früheren und vermutlich auch der aktuellen Bundesregierung: Zwar beinhalte ein Deutschland-Takt „Mustertrassen, die prinzipiell von jedem Eisenbahnverkehrsunternehmen genutzt werden können“, doch sei die dem Deutschland-Takt zugrunde gelegte Betriebsleistung auch mit dem aktuellen Marktmodell möglich (BReg, 2019, S. 3). Grundlage dieser Strategie sind folgende Tatsachen.

Erstens ist das aktuelle Angebot der DB Fern gar nicht so schlecht, es ist auch gar nicht schlecht vertaktet und enthält zugleich HGV-Elemente. Und die DB Fern operiert eigenwirtschaftlich. Hauptprobleme sind die Unpünktlichkeiten und Zugausfälle und mangelnde Verknüpfungen zu nachgeordneten Angeboten des SPFV oder SPNV, und die könnten sich mildern, wenn die infrastrukturellen Voraussetzungen eines Deutschland-Takts geschaffen werden.

Zweitens lässt sich bis auf Weiteres ein noch stärker vertaktetes und verlässliches Taktsystem ohne die infrastrukturellen Ausbauten vermutlich nicht realisieren, und schon gar nicht während der Baumaßnahmen. Ein Gesamtkonzept für den Deutschland-Takt kann daher zunächst gar nicht umgesetzt werden.

Drittens ist gar nicht zu umgehen, dass sich der Deutschland-Takt evolutionär aus dem aktuellen Betriebskonzept der DB Fern entwickeln wird. Das Unternehmen wird von der Politik informell bereits jetzt in diese Richtung gedrängt und ist bemüht, dem nachzukommen, sicherlich auch, um den Takt zu „besetzen“, bevor Wettbewerber eindringen.

Vorschlag 2: Förderung des Wettbewerbs vorbereiten, ohne Einsatz von Steuergeld

Allerdings sollte man sich schon jetzt parallel mit der Förderung des Wettbewerbs im Deutschland-Takt befassen (und selbstverständlich immer Wettbewerb *zum* Deutschland-Takt zulassen). Denn die sukzessive informelle Umsetzung des Deutschland-Takts durch die DB Fern wird an rechtliche Grenzen stoßen. Immer mehr Trassen im Netz und Slots an den Bahnhöfen müssen für „den Takt“ vorreserviert werden, sollen aber faktisch an nur ein Unternehmen gehen. Das wird von der EU zurecht nicht akzeptiert werden.

Eine entstehende Deutschland-Takt-Planbehörde, vielleicht zunächst bei der Bundesnetzagentur, könnte daher bald beginnen, kleine, aber attraktive und in sich stimmige Teile der Deutschland-Takt-Trassen zu

erwerben und potenziellen Wettbewerbern in einer Art Ausschreibung anzubieten, an der sich die DB Fern nicht beteiligen darf. Solche expliziten Förderungen des Wettbewerbs sind in vielen regulierten Sektoren international üblich und würden die genannten Bedenken der EU zerstreuen. Wenn sich Wettbewerber finden, könnte dies im Sinne einer „Ladder of Investment“ den Ansatzpunkt für ihre weitere Entfaltung als Wettbewerber im und zum Deutschland-Takt bilden. Wenn nicht, dann hat man es wenigstens versucht und kann es weiter versuchen.

Dies alles fände unter Beibehaltung des Prinzips der Eigenwirtschaftlichkeit statt, also ohne Subventionen. Bei diesem Prinzip sollte man eigentlich auch bleiben.

Vorschlag 3: Wenn doch Einsatz von Steuergeld, dann möglichst nur im Zubringerbereich über die Länder

Wenn man mit den eigenwirtschaftlich möglichen Angeboten sehr unzufrieden ist, sollte als erstes geprüft werden, ob es ausreichend wäre, subventionierte Verkehre auf den Bereich der sekundären Zubringerverkehre im regionalen Bereich zu beschränken. Diese Verkehre sollten dann im Rahmen einer Aufstockung der Regionalisierungsmittel von den Ländern bestellt werden. Dafür kann auch schnelleres und komfortableres Zugmaterial eingesetzt werden, als für den SPNV üblich ist.

Zwar würde dies nur regionale Shuttle- oder Feederzüge und keine in den primären Fernverkehr durchgebundenen Züge betreffen, aber der Deutschland-Takt zielt ja auch nicht auf durchgebundene Züge, sondern auf gute Umsteigemöglichkeiten ab. Der Vorteil dieser Vorgehensweise bestände darin, dass weiterhin eine klare Trennung zwischen dem Bestellprinzip des Regionalverkehrs und dem Eigenwirtschaftlichkeitsprinzip des primären Fernverkehrs gelten würde. Letzterer würde dann nicht in den Sog der anreiz- und politökonomischen Probleme öffentlich bestellter Verkehre hineingezogen werden.

Vorschlag 4: Wenn doch Einsatz von Steuergeld für primären Fernverkehr, dann mit möglichst klaren Schranken und erst nach Schaffung wettbewerblicher Voraussetzungen

Wenn man hingegen auch die primären Fernverkehrsangebote unter Einsatz von Steuergeldern verbessern möchte, dann sollte dies mit besonderer

Vorsicht geschehen. Bevor subventionierte Verkehre eingerichtet werden, sollten nicht nur die infrastrukturellen, sondern auch die wettbewerblichen Voraussetzungen geschaffen sein. Das heißt, es sollten schon Wettbewerber im Deutschland-Takt etabliert sein. Falls das noch nicht der Fall ist, sollten die ersten Subventionen verwendet werden, um gezielt Wettbewerber zu installieren (Förderung von Werkstätten, Rollmaterial u.a.).

Sodann müssen die politischen Begehrlichkeiten im Zaum gehalten werden. Vielleicht reichen für die Umsetzung einiger Ziele Förderprogramme, die klar definiert und begrenzt sind. Keinesfalls darf jedoch die „Befriedigung der Verkehrsbedürfnisse“ zur Messlatte erhoben werden, der sich die finanzielle Ausstattung anzupassen hätte, sondern umgekehrt müsste immer vom Bund so glaubwürdig wie möglich ein finanzieller Deckel vorgegeben werden.

Vorschlag 5: Richtige Erwartungen an den SPFV

Am überzeugendsten ist, wie gesagt, ein Deckel in Höhe von Null Euro, also die Beibehaltung des Prinzips der Eigenwirtschaftlichkeit, wenigstens für den primären Fernverkehr, da dann gar nicht die Strukturen für Bestellungen geschaffen werden. Dann muss man eventuell auf einige Wünsche an den SPFV verzichten, behält aber ein einigermaßen beherrschbares und funktionierendes System.

Wie können dann die Klimaziele erreicht werden? Gegenfrage: Wie kommt es eigentlich, dass viele Menschen innerhalb Deutschlands billiger fliegen als mit der Bahn fahren können? Mehr Subventionen bekommen die Flieger jedenfalls nicht. Vielleicht funktionieren andere Verkehrssysteme einfach besser als die Bahn. Man darf nicht vergessen, dass die Bahn eine Geldsenke ohnegleichen ist; dies gilt insbesondere für die Eisenbahninfrastruktur, deren Corporate Governance (im ursprünglichen Sinne der Effizienzanreize) wir nicht in den Griff kriegen. Auch mit den Ausbaumaßnahmen, so wichtig einige von ihnen sind, muss man es nicht übertreiben. Sie sind nicht nur in der Anfangsinvestition teuer, sondern ziehen dauerhaft weitere Kosten für Betrieb, Instandhaltung und Ersatzinvestitionen nach sich, die langfristig auch einen Druck auf Erhöhung der Trassenpreise entfalten können. Für das Erreichen der Klimaziele wird Steuergeld dringend im dafür entscheidenden Stromsektor gebraucht. Vielleicht sollte man daher auf die wirtschaftlich funktionierenden Verkehrssysteme bauen, um die große Klima-Transformation auch im Verkehr zu schaffen. Vielleicht müssen nicht massiv zusätzliche Verkehre auf die Bahn verlagert werden, sondern die Menschen könnten in Zukunft auch in Fernbussen,

Autos oder Propellerflugzeugen – alle mit E-Fuels oder Batterien betrieben – ökologisch reisen.

Aber die Bahn wird weiterhin benötigt und sollte die Chance bekommen, auch einen größeren Beitrag leisten zu können. Die Zielvorstellung „Deutschland-Takt“ kann ein geeigneter Rahmen sein, um ihre Systemfähigkeit zusätzlich zu ihrer HGV-Fähigkeit mit Augenmaß weiterzuentwickeln.

Referenzen

- ARGE IGES Institut & IVE (2015). *Machbarkeitsstudie zur Prüfung eines Deutschland-Takts im Schienenverkehr* (Z20/SeV/288.3/1324/LA15). ARGE IGES Institut GmbH & Institut für Verkehrswesen, Eisenbahnbau und -betrieb der Technischen Universität Carolo-Wilhelmina zu Braunschweig (IVE).
- Berschin, F. (2015). Der deutsche Fernverkehr kann deutlich integraler werden. *Eisenbahn-Revue International*, 2015(12), 634–639.
- Berschin, F., Naumann, R., & Nolte, J. (2019). *Der Deutschlandtakt: Bewertung von Organisationsvarianten zur Umsetzung eines flächendeckenden Taktfahrplans* [Gutachten für die Bundtagsfraktion Bündnis 90/Die Grünen]. KCW GmbH. https://www.gruene-bundestag.de/fileadmin/media/gruenebundestag_de/theme_n_az/mobilitaet/pdf/Studie_Deutschlandtakt.pdf.
- BNetzA (2021). *Marktuntersuchung Eisenbahnen 2021: Marktentwicklung 2020 unter den Bedingungen der Covid-19-Pandemie*. Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen.
- Böttger (2021). Marktmodelle zur Organisation des Deutschlandtaktes. Kurzgutachten im Auftrag des Bundesverbands SchienenNahverkehr.
- Böttger, C., & Mitusch, K. (2016). Zur wirtschaftlichen Lage der DB AG. *Zeitschrift Für Verkehrswissenschaft*, 87(3), 211–238.
- Brand, C., & Sieg, G. (2020). The impact of delays on the welfare effects of on-track competition: The case of transfer passengers with operator-tied tickets. *Research in Transport Economics*, 80, 100801.
- Brand, C., & Sieg, G. (2022). An integral interval timetable for long-distance passenger rail services: Time to reconsider targeting on-track competition. *Economics of Transportation*, 32, Article 100285. <https://doi.org/10.1016/j.ecotra.2022.100285>.
- BRat (2016). *Entwurf eines Gesetzes zur Gestaltung des Schienenpersonenfernverkehrs (Schienenpersonenfernverkehrsgesetz—SPFVG)* (Gesetzesantrag der Länder Rheinland-Pfalz, Brandenburg, Bremen, Saarland, BRat-Drucks. 745/16). Bundesrat. https://www.bundesrat.de/SharedDocs/drucksachen/2016/0701-0800/745-16.pdf?__blob=publicationFile&cv=5.

- BReg (2019). *Berücksichtigung der Wettbewerbsneutralität bei der Einführung eines Deutschlandtaktes* (Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Torsten Herbst, Frank Sitta, Oliver Luksic, weiterer Abgeordneter und der Fraktion der FDP, BT-Drucks. 19/11254). Bundesregierung. <https://dserver.bundestag.de/btd/19/112/1911254.pdf>.
- Brenck, A., Mitusch, K., & Dams, J. (2008). *Verbrauchererwartungen an Dienstleistungsqualität im Bahnverkehr* [Studie im Auftrag des Verbraucherzentrale Bundesverband e.V.]. IGES Institut. https://www.iges.com/e6/e34/e10216/e6251/e9716/e9718/attr_objs9724/IGES_VerbrauchererwartungenanDienstleistungsqualitaetimBahnverkehr_Studiefuerdenvzbv_Endfassung_ger.pdf.
- Bundestagsfraktion Bündnis 90/Die Grünen (2019). *Die Eisenbahn zum Rückgrat der Verkehrswende machen* (Antrag, BT-Drucks. 19/7452). Deutscher Bundestag. <https://dserver.bundestag.de/btd/19/074/1907452.pdf>.
- Cascetta, E., & Coppola, P. (2015). New High-Speed Rail Lines and Market Competition: Short-Term Effects on Services and Demand in Italy. *Transportation Research Record: Journal of the Transportation Research Board*, 2475(1), 8–15. <https://doi.org/10.3141/2475-02>.
- Cave, M. (2006). Encouraging infrastructure competition via the ladder of investment. *Telecommunications Policy*, 30(3–4), 223–237. <https://doi.org/10.1016/j.telpol.2005.09.001>.
- Eisenkopf, A., & Knorr, A. (2021). Ein innovativer Regulierungsansatz zur Belebung des Wettbewerbs im Schienenpersonenfernverkehr. *Wirtschaftsdienst*, 101(6), 446–453. <https://doi.org/10.1007/s10273-021-2939-9>.
- FIS (2021). *Potenziale und Herausforderungen des Deutschland-Takts im Schienenpersonenfernverkehr* [Wissenslandkarte]. Forschungs-Informations-System. <https://fis.server.de/servlet/is/538968/?markers=538963>.
- Hendricks, K., Piccione, M., & Tan, G. (1997). Entry and Exit in Hub-Spoke Networks. *The RAND Journal of Economics*, 28(2), 291. <https://doi.org/10.2307/2555806>.
- Hesse, W. (2019). Deutschland-Takt und BMVI-Zielfahrpläne: Chancen, Defizite und Lösungsvorschläge. *Eisenbahn-Revue International*, 2019(7), 386–389.
- Knieps, G. (1996). *Wettbewerb in Netzen: Reformpotentiale in den Sektoren Eisenbahn und Luftverkehr*. Tübingen: Mohr Siebeck.
- Knieps, G. (2007). *Netzökonomie: Grundlagen, Strategien, Wettbewerbspolitik*. Wiesbaden: Gabler.
- Knieps, G. (2008). *Verkehrsinfrastrukturen zwischen Wettbewerb und Regulierung* (Working Paper Diskussionsbeitrag No. 121). Albert-Ludwigs-Universität Freiburg, Institut für Verkehrswissenschaft und Regionalpolitik. <https://www.econstor.eu/handle/10419/32321>.
- Knieps, G., & Brunekreeft, G. (Eds.) (2003). *Zwischen Regulierung und Wettbewerb: Netzsektoren in Deutschland* (2nd ed.). Heidelberg: Physica-Verlag.
- Knieps, G., & Weiß, H.-J. (2009). Regulierung der Eisenbahninfrastruktur: Marktmacht, Interoperabilität und das Defizitproblem. In G. Knieps & H.-J. Weiß (Eds.). *Fallstudien zur Netzökonomie* (S. 139–169). Wiesbaden: Gabler.

- Knieps, G., & Zenhäusern, P. (2007). 'Stepping Stones' and 'Access Holidays': The Fallacies of Regulatory Micro-Management. In P. Baake & R. Borck (Eds.). *Public Economics and Public Choice: Contributions in Honor of Charles B. Blankart* (S. 257–277). Berlin & Heidelberg: Springer.
- Leister, H. (2013). *Deutschland-Takt und Deutschland-Tarif: Relevant für die kommunale Ebene?* Konferenz „Kommunales Infrastruktur-Management“, Berlin. https://www.kim.tu-berlin.de/fileadmin/fg280/veranstaltungen/kim/konferenz_2013/vortraege/noch_www_seminarraum_1-leister_Dt-Takt_Dt-Tarif.pdf.
- Monopolkommission (2017). *Wettbewerbspolitische Baustellen* (Sondergutachten der Monopolkommission gemäß § 78 ERegG, 6. Sektorgutachten Bahn (ehem. SG 76)).
- Monopolkommission (2019). *Mehr Qualität und Wettbewerb auf die Schiene* (Sektorgutachten der Monopolkommission gemäß § 78 ERegG, 7. Sektorgutachten Bahn).
- Monopolkommission (2021). *Bahn 2021: Wettbewerb in den Takt!* (Gutachten der Monopolkommission gemäß § 78 ERegG, 8. Sektorgutachten).
- Redaktion Eurailpress (2018, December 3). Berlin – München: Bahn ist jetzt Verkehrsmittel Nummer Eins. *Eurailpress*. <https://www.eurailpress.de/nachrichten/betrieb-services/detail/news/berlin-muenchen-bahn-ist-jetzt-verkehrsmittel-nummer-eins.html>.
- SMA und Partner, Intraplan Consult, VIA Consulting & Development, & TTS TRIMODE Transport Solutions (2021). *Abschlussbericht zum Zielfahrplan Deutschlandtakt Grundlagen, Konzeptionierung und wirtschaftliche Bewertung* [Entwurf]. Bundesministerium für Verkehr und Digitale Infrastruktur. https://downloads.ctfassets.net/scbs508bajse/7hundjzKsJuMMt31b3tv5a/0e3d5a8e981c3d8068ebc95254e53894/2021-08-31_Abschlussbericht_Deutschlandtakt_1-00_gesamt.pdf.
- Takatsu, T. (2007). The History and Future of High-Speed Railways in Japan. *Japan Railway & Transport Review*, 48, 6–21.
- Tanner, A., & Mitusch, K. (2011). Trassenvermarktung: Auktion versus Listenpreisverfahren. *Internationales Verkehrswesen*, 63(3), 15–19.
- Trainline.com Limited (2020). Trainline: Search, Compare & Buy Cheap Train Tickets. <https://www.thetrainline.com/>.
- Wissenschaftlicher Beirat beim BMVI (2021). Schaffung zukunftsfähiger Grundlagen für einen effizienten Schienengüterverkehr. *Internationales Verkehrswesen*, 73(2), 10–11.

Der Beitrag von (nicht-)ökonomischen Faktoren zur Analyse von Korruption, Steuerhinterziehung, Steuermoral und Schwarzarbeit¹

Friedrich Schneider²

Abstract

In dieser Arbeit werden zunächst einige Ausführungen über die theoretische Modellierung und empirische Evidenz zu wesentlichen Erklärungsfaktoren über Schwarzarbeit und Korruption gegeben. Als klassische Ursachen von Schwarzarbeit gelten die Belastung mit Steuern, Sozialversicherungsbeiträge auf den Produktionsfaktor Arbeit, eine intensive Regulierung, ein unzureichendes und/oder qualitativ schlechtes Angebot an staatlichen Gütern und Dienstleistungen, eine schwache Steuermoral, geringe Rechtssicherheit und nicht funktionierende öffentliche Institutionen, wobei Korruption auch wesentlich von den letzten drei Faktoren verursacht wird. Außerdem werden einige Schätzverfahren und dann empirische Ergebnisse über die Größenordnung von Schwarzarbeit vorgestellt. Anschließend werden aus der Mikroperspektive die (nicht-)ökonomischen Bestimmungsgründe der Steuermoral beispielsweise das Ausmaß der Strafen, die Bildung, das sozialpolitische Engagement erörtert. Um Schwarzarbeit und Korruption einzudämmen, sollte der Staat neben Strafen ein ausreichendes und qualitativ hochstehendes Angebot an öffentlichen Gütern und Dienstleistungen bereitstellen.

-
- 1 Teile dieses Beitrages wurden als schriftliche Fassung meiner Thünen-Vorlesung auf der Jahrestagung des Vereins für Sozialpolitik 2015 zum Thema „Ökonomische Entwicklung, Theorie und Politik“ von Friedrich Schneider „Schwarzarbeit, Steuerhinterziehung und Korruption: Was ökonomische und nicht-ökonomische Faktoren zur Erklärung beitragen“ in den *Perspektiven der Wirtschaftspolitik (PWP)* 16(4), S. 412–425, veröffentlicht.
 - 2 Em. Prof. Dr. Dr.h.c.(mult) Friedrich Schneider, Forschungsinstitut für Bankwesen, Johannes Kepler Universität, Linz, Österreich; friedrich.schneider@jku.at.

JEL-Klassifikation: D78, E26, H2, H11, H26, K42, O5, O17

Keywords:

Schattenwirtschaft, Steuerhinterziehung, Korruption, Schwarzarbeit, ökonomische und nicht-ökonomische Faktoren

1. Einleitung und Problemstellung

Das Ausmaß an Schwarzarbeit und Steuerhinterziehung wird in Deutschland seit einigen Jahren sehr kontrovers diskutiert. Zunächst einmal hierzu einige Fakten zur Steuerhinterziehung aus den Jahren 2019 und 2020: Nach EU-Schätzungen betrug die Steuerhinterziehung bei den *indirekten Steuern* (hier insbesondere die Karussellgeschäfte) in der EU (alle 27 Länder zusammen) circa 50 Mrd. € in 2019 und in Deutschland circa 12 Mrd. €. Für 2020 berechnete der Autor dieses Beitrages, dass die Abgabenhinterziehung bei den *direkten Steuern*, d.h. die üblicherweise bezeichnete „klassische“ Steuerhinterziehung, ein Ausmaß von 16,8 Mrd. € und die, die auf Schwarzarbeit zurückgeht, ein Volumen von 46,5 Mrd. € hatte. Die gesamte Abgabenhinterziehung betrug somit in Deutschland 75,7 Mrd. € oder 10,3 % an den gesamten Einnahmen oder 2,3 % des BIP. Man kann zurecht fragen, ob die Deutschen ein Volk von Steuerhinterziehern und Schwarzarbeitern sind. Im europäischen Vergleich (Durchschnittswerte von 2010 bis 2018) sind sie das nicht, hier liegen sie eindeutig bei der gesamten Abgabenhinterziehung unter dem Durchschnitt von 3,2 % des BIP, den höchsten Wert hat Rumänien mit 6,3 % und den kleinsten die Schweiz mit 1,2 %³. Ebenso kontroverse Ansichten gibt es zur Frage, ob nur hohe Strafen helfen, diese Delikte zu bekämpfen oder ob nicht andere Maßnahmen wie eine anreizorientierte Wirtschaftspolitik in Kombination mit einem qualitativ hochstehenden Angebot von Gütern und Dienstleistungen zur Bekämpfung geeigneter sind.

In diesem Beitrag wird versucht, diese Fragen zu beantworten, wobei auf zwei Themenbereiche eingegangen wird:

- (I) Aus makroökonomischer Sicht werden die Größe der Schwarzarbeit (bzw. Schattenwirtschaft) und das Ausmaß an Korruption geschätzt und es erfolgt eine Analyse ihrer Interaktion (Kapitel 2).

3 Die Quellen dieser Ziffern sind: Poniatowski et al. (2019), eigene Berechnungen sowie Schneider & Buehn (2016).

- (II) Aus mikroökonomischer Sicht wird untersucht, welchen Einfluss auf die Steuermoral der Bürger (und auf ihre tatsächliche Neigung zur Steuerhinterziehung) neben den klassischen ökonomischen Faktoren (Wahrscheinlichkeit der Aufdeckung, Strafe, etc.) nicht-ökonomische wie z.B. soziale Normen und Institutionen haben (Kapitel 3).
- (III) Aus den Erkenntnissen von (I) und (II) werden abschließend einige wirtschaftspolitische Schlussfolgerungen gezogen (Kapitel 4).

2. Die Größenordnung von Schwarzarbeit und Korruption

In diesem Kapitel erfolgen zunächst die Berechnungen über die Größe der Schwarzarbeit/Schattenwirtschaft. Hierbei wird auch die Interaktion mit der Korruption untersucht und welche die wichtigsten Ursachen (oder auch Treiber) für beide sind. Hierbei beschränke ich mich auf das Ausmaß und die Entwicklung der Schwarzarbeit in einigen OECD-Ländern⁴. Anschließend wird untersucht, welchen Einfluss die Korruption auf die Schattenwirtschaft hat und wie groß der volkswirtschaftliche Schaden ist, den die Korruption in der offiziellen Wirtschaft anrichtet. Auf dieser Grundlage wird dann erörtert, welche Maßnahmen zur Reduzierung der Schattenwirtschaft und der Korruption ergriffen werden können.

2.1 Definitionen von Schattenwirtschaft oder Schwarzarbeit

Leider gibt es in der internationalen Forschung keine eindeutige und einheitlich genutzte Definition der Schattenwirtschaft. Eine zentrale Frage ist, welche Aktivitäten fallen überhaupt in die Schattenwirtschaft. Üblicherweise wird folgende Definition verwendet: Tätigkeiten in der Schattenwirtschaft sind die Produktion von Gütern und Dienstleistungen, die legal sind und eine Wertschöpfung darstellen, aber vor den staatlichen Behörden verheimlicht werden. Klassische kriminelle Aktivitäten wie Einbruch, Raub und Drogenhandel fallen üblicherweise nicht in den Bereich der Schattenwirtschaft, die in Deutschland Schwarzarbeit genannt wird. Ebenso wenig gehört die informelle Hauswirtschaft dazu, also sämtliche wirtschaftliche Aktivitäten der Haushalte, die nicht oder nur zum Teil bei der Messung des Bruttoinlandsprodukts (BIP) erfasst werden. Nicht

4 Vergleiche hierzu Schneider (2021), Feld & Schneider (2010), Williams & Schneider (2016), sowie Medina & Schneider (2021).

berücksichtigt werden auch alle „Do-it-yourself“ Aktivitäten sowie die Nachbarschaftshilfe, bei der einander zumeist unentgeltlich geholfen wird. Da sich die genannten Tätigkeiten überlappen, können je nach Definition und Messverfahren die empirischen Annäherungen an das Ausmaß der Schattenwirtschaft sehr unterschiedlich ausfallen.

2.2. Ursachen und Indikatoren von Schattenwirtschaft und Korruption

Es gibt eine Vielzahl von Treibern oder Ursachen für die Schattenwirtschaft und die Korruption. Theoretisch lassen sich insbesondere die folgenden acht Faktoren begründen⁵:

- **Steuern und Sozialversicherungsbeiträge** – je höher die Abgabenlast liegt, desto mehr wird schwarzgearbeitet und desto größer ist damit die Schattenwirtschaft.
- **Regulierung der Wirtschaft** – je mehr der Staat reguliert, desto mehr wird schwarzgearbeitet, desto größer fällt also die Schattenwirtschaft aus.
- **Mangelndes Angebot an staatlichen Gütern und Dienstleistungen** – je weniger Güter und Dienstleistungen der Staat selbst anbietet, desto mehr erarbeiten sich die Menschen im Schatten.
- **Mangelnde Steuermoral** – je geringer die Steuermoral der Bürger ausgeprägt ist, desto mehr arbeiten sie schwarz und umso größer ist die Schattenwirtschaft.
- **Mangelnde Qualität der öffentlichen Institutionen** – Je schlechter die öffentlichen Institutionen funktionieren, desto mehr Arbeit wandert in die informelle Welt ab, umso größer ist also die Schattenwirtschaft.
- **Korruption** – je mehr bestochen werden muss und wird, desto mehr arbeiten die Menschen schwarz.
- **Geringe Rechtssicherheit** – je schlechter sich private Eigentumsrechte durchsetzen können, desto größer fallen Schwarzarbeit und Korruption aus.
- **Geringe politische Partizipation** – je weniger die Bürger ein direktes Mitspracherecht haben, desto mehr arbeiten sie schwarz bzw. umso größer ist die Schattenwirtschaft. Oder umgekehrt argumentiert: Wenn

5 Auf eine ausführliche Erörterung sowie Herleitung wird aus Platzgründen verzichtet. Vgl. hierzu Schneider & Enste (2000), Schneider (2015, 2021) und Feld & Schneider (2010).

die Bürger, ein starkes direktes Mitspracherecht haben, zum Beispiel in einem Föderalsystem und/oder mit direkter Demokratie, dann reduziert dieses Element der politischen Kultur den Umfang der Schattenwirtschaft.

Hauptindikatoren, in denen sich die Größe der Schattenwirtschaft widerspiegelt, sind:

- **die im Umlauf befindliche Bargeldmenge** (das Verhältnis von M_0 zu M_1) – denn je mehr schwarzgearbeitet wird bzw. je größer die Schattenwirtschaft ist, desto mehr Bargeld kommt zum Einsatz.
- **das BIP-Wachstum** – denn je mehr schwarzgearbeitet wird bzw. je größer die Schattenwirtschaft ist, desto schwächer wächst das Bruttoinlandsprodukt in der offiziellen Wirtschaft.
- **die Erwerbsquote** – denn je mehr schwarzgearbeitet wird bzw. je größer die Schattenwirtschaft ist, desto geringer ist der Anteil der Erwerbstätigen in der offiziellen Wirtschaft.

Die Interaktion von Schwarzarbeit und Korruption ist allerdings in der Wissenschaft sowohl theoretisch als auch empirisch umstritten. Es gibt zwei traditionelle Hypothesen: Schwarzarbeit und Korruption sind entweder Substitute (z.B. Rose-Ackermann, 1997) oder Komplemente (z.B. Choi & Thum, 2005; Friedman et al., 2000). Dazu kommt die neuere, differenzierte Hypothese von Dreher und Schneider (2009), nach welcher Schwarzarbeit und Korruption in hochentwickelten (OECD) Ländern Substitute, in Entwicklungsländern jedoch Komplemente sind. Die empirische Evidenz über die Interaktion von Schwarzarbeit und Korruption ist allerdings nicht eindeutig.⁶

Die wesentlichen Ursachen/Treiber für Korruption sind:

- Geringe Qualität sowie Effektivität von öffentlichen und rechtlichen Institutionen – je schlechter das politische System und die Durchsetzungsfähigkeit des Rechtsstaats sind, desto höher ist die Korruption.
- Wenig soziales und kulturelles Kapital – je niedriger beispielsweise das allgemeine Bildungsniveau ist, desto mehr Korruption gibt es.

6 Vergleiche hierzu insbesondere die Arbeit von Dreher & Schneider (2009).

- Wenig Freiraum für die Wirtschaft – je größer der Staatssektor ist, bzw. je stärker der Staat in die Wirtschaft eingreift, desto mehr Korruption lässt sich beobachten.

Indikatoren für das Ausmaß an Korruption sind:

- BIP pro Kopf – denn je stärker die Korruption ist, desto schwächer entwickelt sich die offizielle Wirtschaft⁷.
- Ein Bestechungs-Index, der meistens in Panelstudien abgeleitet wird, etwa aufgrund der Frage: „Was schätzen Sie, wie üblich es in Ihrer Branche ist, dass Firmen nicht-dokumentierte Zahlungen vornehmen oder Bestechungsgelder zahlen?“ – denn je stärker die Korruption ist, desto höher fällt der Indexwert aus.
- Unabhängigkeit der Justiz – je stärker die Korruption ist, desto schlechter ist es um die Unabhängigkeit der Justiz bestellt.

2.3 Schätzverfahren und empirische Ergebnisse

Zur Messung bzw. Schätzung der Schattenwirtschaft und Korruption gibt es mehrere Schätzverfahren⁸. In diesem Beitrag gehe ich nur auf ein einziges Schätzverfahren ein, und zwar auf das MIMIC-Schätzverfahren (Multiple Indicators Multiple Causes). Im MIMIC-Verfahren ist die Schattenwirtschaft (und/oder auch die Korruption) eine nicht-beobachtbare (latente) Variable. Die Beziehung zwischen der nicht-beobachtbaren Variable und deren Ursachen wird in dem folgenden strukturellen Modell abgebildet:

$$\eta = \gamma x + \varsigma \quad (1)$$

wobei:

η : nicht beobachtbare Variable (Schattenwirtschaft oder Korruption),

x : q-Vektor von Ursachen im strukturellen Modell, und

γ : q-Vektor von Koeffizienten der Ursachen im strukturellen Modell.

-
- 7 Hier wird die Annahme gemacht, dass zumeist die „Sand-in-the-wheel“-Hypothese dominiert. Sie besagt, dass Korruption die wirtschaftliche Entwicklung bremst und zu geringerem Wohlstand führt. Im Gegensatz dazu gibt es aber auch die „Grease“-Hypothese, nach der Korruption als Schmiermittel zugunsten einer Steigerung der wirtschaftlichen Effizienz wirkt. Vergleiche hierzu Meon & Weill (2010).
- 8 Da es hierzu sehr viel Literatur gibt, wird dieses Thema nicht weiter erörtert. Vergleiche hierzu insbesondere Schneider & Enste (2000), Feld & Schneider (2010), Medina & Schneider (2021), Schneider (2021).

Ebenso wird die Beziehung zwischen der nicht-beobachtbaren Variable und deren Indikatoren in einem Messungsmodell abgebildet:

$$y = \lambda\eta + \varepsilon \quad (2)$$

wobei:

y: p-Vektor von Indikatoren im Messungsmodell,

λ : p-Vektor von Koeffizienten der Indikatoren im Messungsmodell, und

ζ, ε : Fehlerterme im strukturellen und Messungsmodell.

Im MIMIC-Schätzverfahren werden die folgenden strukturellen und Indikatoren-Gleichungen spezifiziert. Dies erfolgt am Beispiel der Schattenwirtschaft:

Strukturelle Gleichung, die die Ursachen (Treiber) für die Schattenwirtschaft spezifiziert:

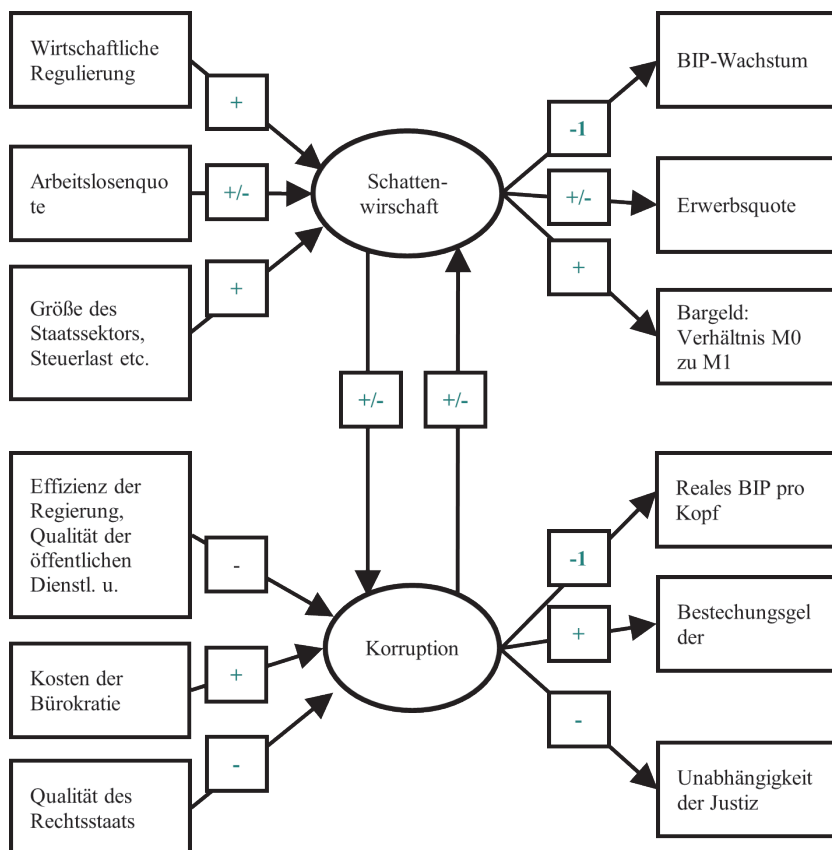
$$[\text{Schattenwirtschaft}] = [\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7, \gamma_8] * \begin{bmatrix} \text{Anteil direkter Steuern} \\ \text{Anteil indirekter Steuern} \\ \text{Anteil Sozialversicherungsbeiträge} \\ \text{Regulierungslast} \\ \text{Qualität staatl. Institutionen} \\ \text{Steuermoral} \\ \text{Arbeitslosenquote} \\ \text{BIP pro Kopf} \end{bmatrix} + \zeta$$

und die folgende Messgleichung, in der sich die Schattenwirtschaftsaktivitäten widerspiegeln:

$$\begin{bmatrix} \text{Erwerbsquote} \\ \text{Veränd. Bargeldmenge} \\ \text{BIP – Wachstum} \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} * \text{Schattenwirtschaft} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

In Abbildung 1 ist ein MIMIC-Modell mit zwei latenten Variablen dargestellt, der Schattenwirtschaft und der Korruption. Im linken Teil der Grafik ist die Wirkungsweise der Einflussfaktoren und im rechten Teil die der Indikatoren aufgeführt. Die Arbeitslosenquote, die wirtschaftliche Regulierung und die Größe des Staatssektors wirken beispielsweise erhöhend auf die Schattenwirtschaft, wie das positive Vorzeichen anzeigt. Die Aktivitäten in der Schattenwirtschaft spiegeln sich in einem niedrigeren BIP-Wachstum, in einer tieferen offiziellen Erwerbsquote und in einem Anstieg der Bargeldnachfrage wider. Die Kosten der Bürokratie erhöhen die Korruption und je besser der Rechtsstaat ausgeprägt ist, desto geringer ist die Korruption. Eine hohe Korruption schlägt sich in einem tieferen BIP nieder.

Abbildung 1: MIMIC-Modell für die Schattenwirtschaft und Korruption



Quelle: Buehn & Schneider (2012, S. 185).

In Tabelle 1 sind die Ergebnisse der MIMIC-Schätzungen für die Schattenwirtschaft und Korruption über 51 Länder und über die Periode 2000 bis 2005 aufgeführt.⁹ Je intensiver die Regulierung der Wirtschaft, je höher die Arbeitslosenquote und die Steuerbelastung sind, desto größer ist die Schattenwirtschaft; diese Einflussfaktoren sind statistisch signifikant. Die Korruption wird durch die Regulierung der Bürokratie erhöht, und sie

⁹ Diese Schätzungen sind von Buehn & Schneider (2012, S. 186f.) entnommen. Auf eine detaillierte Interpretation unter Berücksichtigung des ceteris paribus Argumentes wird hier verzichtet.

sinkt durch gut funktionierende staatliche Institutionen (Ursachen-Variable: Effektivität des Staates) und eine größere marktwirtschaftliche Freiheit. Die Indikatoren zeigen, dass die Erwerbsquote (in % der erwerbsfähigen Bevölkerung) und das Bargeld (Verhältnis von M0 zu M1) die wichtigsten Indikatoren für den Umfang der Schattenwirtschaft sind. Im Fall der Korruption sind dies die Bestechungsgelder. In der MIMIC-Schätzung wird auch die Interaktion von Schattenwirtschaft und Korruption untersucht. Beide wirken verstärkend aufeinander ein und sind mithin komplementär, wobei die Schattenwirtschaft die Korruption etwas stärker beeinflusst als umgekehrt.

Tabelle 1: Schätzergebnisse für Schattenwirtschaft und Korruption in 51 Ländern 2000–2005

Ursachen	Schattenwirtschaft	Korruption
Wirtschaftliche Regulierung (Index)	0,18** (2,00)	
Arbeitslosenquote (in %)	0,19** (1,98)	
Größe des Staates (in % des BIP)	0,09 (1,16)	
Steuerbelastung (gesamte Steuereinnahmen in % des BIP)	0,16** (1,98)	
Effektivität des Staates (Index)		-0,22*** (3,13)
Wirtschaftliche Freiheit (Index)		-0,15*** (2,48)
Kosten der Bürokratie (Index)		0,42*** (5,15)
Qualität des Rechtsstaates (Index)		-0,01 (0,10)
Indikatoren		
BIP-Wachstum (in %)	-0,51	
Erwerbsquote (in % der erwerbstätigen Bevölkerung)	-0,41*** (4,15)	

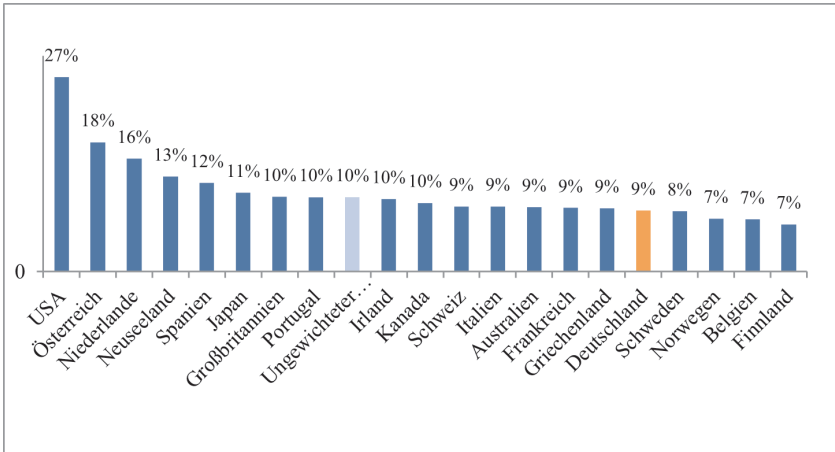
Bargeld (Verhältnis von M0 zu M1)	0,31*** (3,33)	
Reales BIP pro Kopf		-0,78
Bestechungsgelder (Index)		0,15* (1,73)
Unabhängigkeit der Justiz (Index)		-0,06 (0,73)
Beziehung zwischen den nicht-beobachtbaren Variablen		
Schattenwirtschaft → Korruption	0,68*** (4,23)	
Korruption → Schattenwirtschaft	0,42*** (2,64)	
Beobachtungen	168	
Chi-Quadrat (p-Wert)	97,23 (0,11)	
RMSEA	0,04	
AGFI	0,90	

Anmerkung: Die absoluten z-Statistiken sind hier in Klammern angegeben. Die Symbole *, **, *** zeigen die Signifikanz am Niveau von 10 %, 5 % bzw. 1 %. Quelle: Buehn & Schneider (2012, S. 186f.).

In den Abbildungen 2 und 3 sind die neuesten Ergebnisse der Schattenwirtschaft in 20 OECD Ländern aufgeführt. Die Berechnungen wurden mit dem MIMIC-Verfahren durchgeführt¹⁰. Durch die globale Coronapandemie und die dadurch ausgelöste schwere Rezession ist die Schattenwirtschaft sowohl im Jahr 2020 als auch in 2021 (wenn auch im geringeren Umfang) in allen 20 OECD Ländern stark gestiegen. In Abbildung 2 ist dieser prägnante Anstieg von 2019 auf 2020 aufgeführt. Die Hauptursachen hierfür sind der markante Rückgang des BIP und der starke Anstieg der Arbeitslosigkeit und der Kurzarbeit. In den USA hat die Schattenwirtschaft von 2019 auf 2020 proportional am stärksten zugenommen, nämlich um 27 %; allerdings haben die USA mit 5,9 % vom BIP neben der Schweiz auch den tiefsten Wert an Schattenwirtschaft in 2019.

10 Auf einer Darstellung der Schätzergebnisse wird aus Platzgründen verzichtet. Vergleiche hierzu Schneider & Boockmann (2021).

Abbildung 2: Anstieg der Schattenwirtschaft in ausgewählten Industrieländern 2019 bis 2020

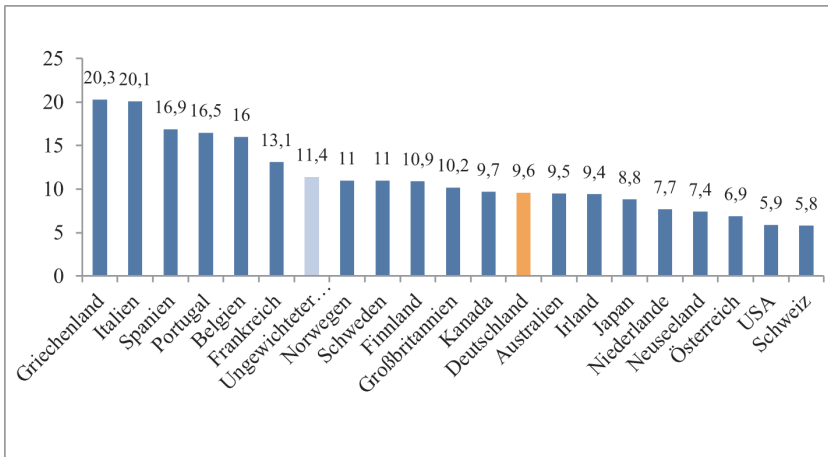


Erläuterung: Aufgeführt ist die prozentuale Zunahme des Verhältnisses der Schattenwirtschaft zum offiziellen BIP von 2019 auf 2020. Quelle: Schneider & Boockmann (2021).

In Abbildung 3 sind die Ergebnisse zur Größe der Schattenwirtschaft in 20 europäischen Ländern für das Jahr 2021 dargestellt. Dies sind simulierte bzw. zum Teil prognostizierte Werte, die sich, je nachdem wie stark die wirtschaftliche Erholung in 2021 noch ausfällt, noch ändern können. Deutschland hat eine Schattenwirtschaft in der Größe von 9,6 % des BIP, Österreich von 6,9 % und die Schweiz von 5,8 %. Die durchschnittliche anteilige Größe der Schattenwirtschaft für die 20 Länder beträgt 11,4 % des BIP, wobei die südeuropäischen Länder Griechenland, Italien und Spanien mit Anteilen von 20,3 %, 20,1 % und 16,9 % die größten Schattenwirtschaften aufweisen.¹¹

11 Auf eine ausführliche Diskussion dieser Ergebnisse wird hier verzichtet. Siehe hierzu Schneider & Boockmann (2021).

Abbildung 3: Niveau der Schattenwirtschaft in ausgewählten Industrieländern 2021 (im Verhältnis zum BIP)



Erläuterung: Aufgeführt ist das Verhältnis der Schattenwirtschaft zum offiziellen BIP in Prozent für das Jahr 2021. Quelle: Schneider & Boockmann (2021).

Weitere Studien, die die Ursachen für Schattenwirtschaft und Korruption analysieren¹², bestätigen, dass eine stärkere Steuerbelastung und eine intensivere Regulierung das Ausmaß der Schattenwirtschaft erhöhen. Ebenso zeigen sie, dass eine größere Schattenwirtschaft den Bargeldumlauf erhöht. Nur in einigen Untersuchungen kann bestätigt werden, dass sich das Wachstum der offiziellen Wirtschaft mit zunehmender Größe der Schattenwirtschaft verlangsamt. Was die Korruption betrifft, kann nur teilweise bestätigt werden, dass sie umso mehr Raum gewinnt, je schlechter die staatlichen Institutionen, d.h. die Durchsetzungsfähigkeit des Rechtsstaates sind. Dass auch ein geringes Bildungsniveau die Korruption treibt, lässt sich empirisch nicht bestätigen. Eindeutig bestätigt wird hingegen, dass die Korruption steigt, wenn im Land wenig wirtschaftliche Freiheit herrscht. Dass ein Zuwachs der Korruption die wirtschaftliche Entwicklung bremst, kann empirisch nur für manche Länder (insbesondere Entwicklungsländer) bestätigt werden.

12 Hier werden nur die wichtigsten Ergebnisse zusammengefasst; in Schneider (2021), Medina & Schneider (2021), Buehn & Schneider (2012) sowie Feld & Schneider (2010) erfolgt eine ausführliche Diskussion.

2.4 Kritik an empirischen Schätzungen

Es sei ausdrücklich erwähnt, dass alle Schätzverfahren zur Ermittlung der Größe der Schattenwirtschaft und der Korruption Schwächen haben und deswegen auch häufig kritisiert werden.¹³ Beispielhaft sind folgende beiden Kritikpunkte erwähnt: (i) Beim oben angewendeten MIMIV-Schätzverfahren gibt es keine von der Theorie eindeutig gestützte „Regel“ bzw. eine klare Trennung, welches die Ursachen- und welches die Indikator-Variablen sind. (ii) Auch liefert das MIMIC-Verfahren nur relative Schätzungen der Größe der Schattenwirtschaft und der Korruption. Wegen der notwendigen Kalibrierung der Schätzergebnisse mit einem anderen Schätzverfahren (häufig wird hier die Bargeldmethode verwendet) kommt es zu einer großen Bandbreite an Schätzergebnissen.

2.5 Wirtschaftspolitische Maßnahmen

Aus der Kenntnis über die Ursachen von Schwarzarbeit und Korruption werden wirtschaftspolitische Maßnahmen zu ihrer Reduktion abgeleitet. Aus einer Vielzahl von Maßnahmen, mit denen die Politik versuchen kann, Aktivitäten in der Schattenwirtschaft und Korruption einzudämmen, seien die folgenden erwähnt. Die Schattenwirtschaft sinkt, wenn die Abgabenlast auf die Lohnnebenkosten verringert wird. Man kann darüber hinaus entweder die Umsatzsteuer auf arbeitsintensive Dienstleistungen rückvergüten, oder Haushaltsinvestitionen von der Steuer absetzbar machen. Gegen die Korruption kann man unter anderem mit Sanktionen vorgehen, zum Beispiel indem man korrupte Unternehmen für eine bestimmte Zeit, beispielsweise für 3 bis 5 Jahre, für öffentliche Aufträge sperrt – oder indem man staatlichen Bediensteten und Politikern, denen Bestechlichkeit nachgewiesen worden ist, die Pensionsansprüche streicht. Auch eine aktive Unterstützung und ein stärkerer rechtlicher Schutz von „Whistleblowern“ sind hilfreich.

13 Auch hier werden aus Platzgründen nur einige Kritikpunkte genannt. Vgl. hierzu insbesondere Schneider & Enste (2000), Feld & Schneider (2010), Schneider (2021).

3. *Steuermoral und ihre Ursachen aus der Mikroperspektive*

Warum zahlen die meisten Individuen ehrlich ihre Steuern? Aus Sicht der Public Choice Theorie, in der eigennütziges Verhalten dem Steuerzahler unterstellt wird, ist die sehr hohe Anzahl der ehrlichen Steuerzahler nicht erklärbar. In den meisten Ländern ist die Kontrollintensität sehr gering, die Strafen sind zumindest beim ersten Mal „erwischt“ werden nicht hoch und häufig verfolgen die Behörden fiskalische Strafverfahren nicht sonderlich energisch. Somit liegt die Frage auf der Hand, warum nicht mehr Bürger versuchen, ihre Steuerlast jenseits der Legalität ein wenig zu reduzieren?

3.1 *Die Ablehnung des Steuerbetruges durch die Bürger*

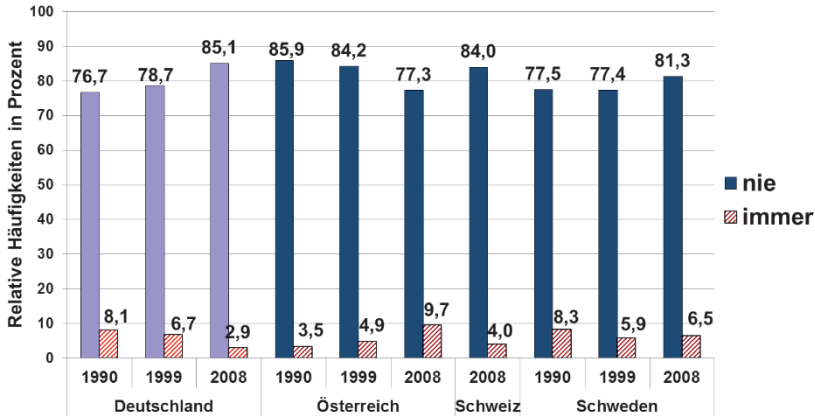
In Abbildung 4 sind die Umfrageergebnisse des European Value Survey (EVS) bei der Frage „Halten sie Steuerhinterziehung für gerechtfertigt“ für die Jahre 1990, 1999 und 2008 und für Deutschland, Österreich, Schweden und die Schweiz aufgeführt. Die Ergebnisse zeigen eindeutig, dass in diesen vier Ländern die Neigung der Bürger zur Steuerhinterziehung ausgesprochen gering ist. Wenn man sie fragt, ob sie Steuerhinterziehung für gerechtfertigt halten, antworten in Deutschland, in Österreich, in der Schweiz und in Schweden jedenfalls weit mehr als 75 % der Bürger mit „nie“.

In Abbildung 5 sind die Ergebnisse der Umfrage „Tolerieren sie ungerechtfertigten Bezug staatlicher Leistungen?“ für die gleichen Jahre und vier Länder aufgeführt. Hier fällt das Ergebnis noch eindeutiger aus. In Deutschland antworten weit mehr als 80 % darauf „nie“. In Österreich ist dieser Anteil erst 2008 unter diese Schwelle gesunken, während er 1990 bei 92,1 % lag. In der Schweiz sind es fast 92 % und in Schweden knapp 88 %.

Was veranlasst die Steuerzahler zu einem derart hohen Bekenntnis zur Ehrlichkeit? Aus ökonomischer Perspektive wäre es denkbar, dass sie die Wahrscheinlichkeit, mit ihrem illegalen Tun aufzufliegen, schlicht falsch einschätzen. Die in der Entscheidung angenommene subjektive Wahrscheinlichkeit der Entdeckung von Steuer- und Sozialleistungsbetrug durch die Behörden wäre damit also wesentlich höher als die tatsächliche Wahrscheinlichkeit. Hierzu gibt es in der Tat einige Evidenzen.

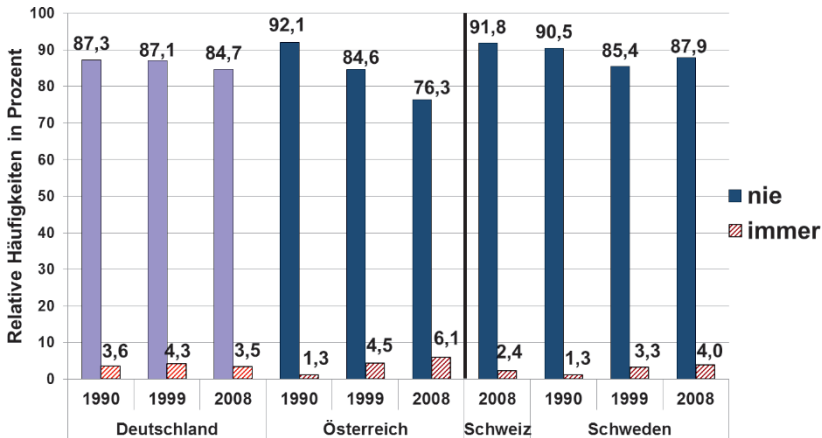
Es ist allerdings auch möglich, dass die Menschen nur bedingt ökonomisch rational abwägen, sondern eine weitgehend unabhängig von ökonomischen Erwägungen eher von allgemeinen moralischen Normen (z.B. Re-

Abbildung 4: „Halten Sie Steuerhinterziehung für gerechtfertigt?“ Ergebnisse des European Value Survey (EVS) 1990, 1999 und 2008.



Quelle: EVS (2015).

Abbildung 5: „Tolerieren Sie den ungerechtfertigten Bezug staatlicher Leistungen?“ Ergebnisse des European Value Survey (EVS) 1990, 1999 und 2008.



Quelle: EVS (2015)

ligion, Patriotismus, Bildung, ehrenamtliche Tätigkeiten) bestimmte Steuermoral haben. Dieses Argument haben schon Allingham und Sandmo in ihrer bahnbrechenden Arbeit im Jahr 1972 vorgebracht. Sie und viele

andere Autoren sind der Ansicht, dass fundamentale soziale Normen oder konsensuale Vorstellungen über Bürgerpflichten die Steuermoral prägen.¹⁴ Wenn das Konzept der Steuermoral erklären soll, warum das Bekenntnis zur Steuerehrlichkeit so deutlich ausfällt, wie es die Umfragen zeigen, dann werden Untersuchungen benötigt, welche Faktoren die Steuermoral prägen.

3.2 Ein fiktiver Vertrag zwischen Staat und Steuerzahler

Wie bereits erwähnt, hängt die individuelle Steuermoral von einer Vielzahl verschiedener Einflüsse ab. Beispielsweise hängt die Bereitschaft Steuern zu zahlen, auch wesentlich davon ab, wie viele andere Individuen derjenige kennt, die ebenso bereitwillig ihre Steuern und Abgaben zahlen. Eine öffentliche Diskussion über eine (angeblich) sinkende Steuermoral kann den ehrlichen Steuerzahlern den Eindruck vermitteln, sie seien die letzten und die einzigen („Dummen“), die Steuern zahlten und das kann dazu führen, dass ihre Bereitschaft Steuern zu zahlen abnimmt. Im Gegenzug stellt der Staat mit den Einnahmen aus Steuern und anderen Abgaben öffentliche Güter und Dienstleistungen bereit. Wenn die Belastung mit Steuern und Abgaben als zu hoch und ungerecht empfunden wird und den individuell empfundenen Wert der genutzten Güter und Dienstleistungen in eklatanter Weise übersteigt, erleichtert dies die Steuerhinterziehung. Diese Einschätzung dürfte umso drastischer sein, je unzufriedener die Bürger mit dem staatlichen Angebot an Gütern und Dienstleistungen sind.

Üblicherweise nehmen die Bürger ihre Steuerzahlungen als Beitrag zum „Bonum commune“ wahr. Sie sind deshalb auch durchaus bereit, ihr Einkommen gegenüber dem Fiskus ehrlich zu deklarieren, selbst wenn sie nicht in jedem Fall eine gleichwertige Menge oder Qualität an öffentlichen Gütern für ihre Steuerzahlungen erhalten. Viele Menschen sind Nettozahler für den Wohlfahrtsstaat und nehmen die steuerliche Umverteilung sogar bewusst in Kauf, weil sie in einem sicheren und sozial ausgeglichenen Land leben möchten. Und schließlich spielt auch die Art und Weise, wie die Behörden die Steuerzahler behandeln – insbesondere bei Steuerprüfungen – eine wichtige Rolle. Die Steuerehrlichkeit als auch die Steuermoral

14 Vgl. hierzu beispielsweise die Arbeiten von Feld & Frey (2004, 2007), Halla & Schneider (2005, 2014), Schneider et al. (2010) und Torgler & Schneider (2007, 2009).

nimmt zu, wenn sich die Steuerzahler als Kunden und mündige Bürger fühlen und nicht als Untertanen, die bloß zu gehorchen haben.

Feld und Frey (2007) modellieren diese Einflüsse in der Weise, dass daraus ein Vertrag zwischen den Steuerzahlern und den Steuerbehörden bzw. der Regierung entsteht, der an einige Bedingungen geknüpft ist. Dieser Vertrag stellt somit ein Tauschgeschäft dar: Der Staat bekommt die Steuereinnahmen der Bürger, und diese im Gegenzug Güter und Dienstleistungen vom Staat. Damit der Vertrag Bestand hat, müssen die Vertragspartner, also Steuerzahler und Steuerbehörden, einander mit Respekt und Ehrlichkeit begegnen. Den Steuerzahler können allerdings drei Dinge davon abhalten, den Vertrag zu erfüllen: dass ihn der Fiskus schließlich doch nicht als ebenbürtigen Partner behandelt, sondern als Untertan; dass das staatliche Angebot an Gütern und Dienstleistungen unzureichend oder schlecht ist; oder dass die staatlichen Institutionen durch die Möglichkeit der Korruption versagen.

3.3 Ein empirischer Erklärungsversuch der Steuer- und Sozialstaatsmoral

Abschließend erfolgt eine knappe Darstellung einer empirischen Untersuchung, aus welchen Faktoren sich die Steuer- und Sozialstaatsmoral unter ceteris paribus Bedingungen zusammensetzt. In einer Fallstudie haben Halla und Schneider (2005, 2014) die Daten der österreichischen Befragten des European Value Survey (EVS) für 1990 und 1999 ausgewertet. Sie sind zu dem Ergebnis gekommen, dass sowohl ökonomische als auch nicht-ökonomische Faktoren die Steuermoral der Bürger beeinflussen; das Gleiche gilt für die Sozialstaatsmoral. Darunter verstehen sie eine grundsätzliche normative Haltung zu den Leistungen des Sozialstaats: Wer eine hohe Sozialstaatsmoral hat, hält es nur dann für gerechtfertigt, sozialstaatliche Leistungen in Anspruch zu nehmen, wenn sie einem legal tatsächlich zustehen. Ziel der Arbeiten von Halla und Schneider ist es die abhängigen (d.h. zu erklärenden) Variablen Steuermoral und Sozialstaatsmoral so gut wie möglich mit unabhängigen Variablen zu erfassen. Die beiden abhängigen Variablen gehen auf zwei Aussagen von Befragungen des EVS zurück, die sie gemäß einer 10-Punkte-Skala (1=immer, 10=nie) gewichten:

- „Steuern zu hinterziehen, wenn man die Möglichkeit dazu hat, ... ist immer gerechtfertigt, nie gerechtfertigt, oder etwas dazwischen“ (Steuermoral), und

- „Staatliche Sozialleistungen zu beziehen, auf die man eigentlich keinen Anspruch hat, ... ist immer gerechtfertigt, nie gerechtfertigt, oder etwas dazwischen“ (Sozialstaatsmoral).

Insgesamt stehen hierfür 1.887 Beobachtungen zur Verfügung, 835 für 1990 und 1.052 für 1999. Daraus werden die folgenden beiden Schätzgleichungen gebildet:

$$\textbf{Steuer}moral = \alpha_1 + \beta_{11} * \textbf{Sozial}staatsmoral + \beta_{12} * \textbf{Einkommen} + \Gamma_1 * X_1 + \varepsilon_1$$

$$\textbf{Sozial}staatsmoral = \alpha_2 + \beta_{21} * \textbf{Steuer}moral + \beta_{22} * \textbf{Einkommen} + \Gamma_2 * X_2 + \varepsilon_2$$

Wenn die in den Umfragen zum Ausdruck kommende Steuermoral bzw. Sozialstaatsmoral Einfluss auf das tatsächliche Verhalten hat, ist allerdings das Einkommen ebenfalls endogen, d.h. Steuermoral, Sozialstaatsmoral und Einkommen beeinflussen sich gegenseitig, da Simultanität besteht. Halla und Schneider (2005, 2014) machen das Einkommen zu einer Instrumental-Variable, indem sie es mit dem Geschlecht und der Beschäftigungsart des Hauptverdieners sowie mit dem sozioökonomischen Status des Haushaltes zu erklären versuchen.

In Tabelle 2 sind die Ergebnisse der „Two-stage least square“-Schätzungen (2SLS) für die Steuer- und Sozialstaatsmoral aufgeführt¹⁵: Die Sozialstaatsmoral hat einen positiven Einfluss auf die Steuermoral, und umgekehrt wirkt die Steuermoral auch positiv auf die Sozialstaatsmoral. Einkommen und Schulbildung hingegen üben einen (keinen) negativen Einfluss auf die Steuermoral (Sozialstaatsmoral) aus: Je mehr Einkommen jemand erzielt und je mehr Schuljahre er absolviert hat, desto schlechter ist die Steuermoral, ceteris paribus. Hingegen wirkt sich Religiosität positiv auf die Steuermoral aus: je religiöser der Mensch, desto besser seine Steuermoral. Weiterhin verschlechtert starkes Misstrauen gegenüber dem Rechtssystem die Steuermoral. Die Sozialstaatsmoral wird stark vom Alter beeinflusst, je älter jemand wird, desto höher ist seine Sozialstaatsmoral, ceteris paribus. Sie hängt auch von der Erwerbstätigkeit ab, ist jemand beschäftigt, ist seine Sozialstaatsmoral höher. Auch die Bereitschaft zu freiwilligen Tätigkeiten, steigert die Sozialstaatsmoral.

15 Auf eine ausführliche Darstellung der ökonometrischen Vorgehensweise als auch auf eine ausführliche Interpretation der Schätzergebnisse wird hier verzichtet; der interessierte Leser sei auf die beiden Veröffentlichungen von Halla und Schneider verwiesen.

Tabelle 2: 2SLS-Schätzung von Steuer- und Sozialstaatsmoral

Variable	Steuermoral	Sozialstaatsmoral
Steuermoral (Index)	-	0,463** (0,220)
Sozialstaatsmoral (Index)	0,607** (0,282)	-
Einkommen (€ 100 p. M.)	-0,024* (0,013)	0,002 (0,013)
Alter (Jahre)	-0,004 (0,005)	0,011*** (0,003)
Weiblich (=1, Dummy-Variable)	0,003 (0,094)	0,122 (0,077)
Verheiratet (=1, Dummy-Variable)	0,182* (0,105)	-0,043 (0,109)
Bildung (Schuljahre)	-0,022* (0,013)	0,001 (0,013)
Beschäftigt (=1, Dummy-Variable)	-0,083 (0,130)	0,226** (0,096)
1999 (=1, Dummy-/Shift-Variable)	0,141 (0,180)	-0,480*** (0,090)
Religiosität (Index)	0,136*** (0,053)	-0,006 (0,061)
Patriotismus (Index, 1=wenig Patriotismus; 4=großer Patriotismus)	0,157 (0,101)	0,132 (0,093)
Misstrauen gegenüber dem Rechtssystem (Index, 1=kein Misstrauen, 4=sehr großes Misstrauen)	-0,137** (0,054)	-
Kinder (Anzahl)	-	0,073** (0,032)
Freiwillige Tätigkeiten (=1, Dummy-Variable)	-	0,136* (0,083)
Konstante	3,410 (2,151)	3,958** (1,671)
Beobachtungen	1.887	1.887
Basmann-Statistik ^a	0,237	0,584
Sargan-Statistik ^b	0,232	0,580

Anmerkung: Standardfehler in Klammern. *, ** und *** zeigen die statistische Signifikanz auf dem Niveau von 10 %, 5 % und 1 %. Dummies für die neun österreichischen Bundesländer sind in jeder Regression berücksichtigt. Basisgruppe ist Vorarlberg.

^a P-Wert von Basmann's Test zur Überidentifizierung der Restriktionen aller Instrumente.

^b P-Wert von Sargan's Test zur Überidentifizierung der Restriktionen aller Instrumente.

Die Anwendung von Probit-Schätzungen auf beiden Stufen der Schätzung der Steuer- und Sozialstaatsmoral-Gleichungen liefert qualitativ ähnliche Ergebnisse.

Quelle: Halla & Schneider (2005, 2014).

Da Steuermoral, Sozialstaatsmoral und Einkommen einander wechselseitig beeinflussen und mithin Endogenität in beide Richtungen besteht, wird abschließend noch umgekehrt der Einfluss der Steuer- und Sozialstaatsmoral auf das Einkommen untersucht:

$$\text{Einkommen} = \alpha_3 + \beta_{31} * \text{Steuermoral} + \beta_{32} * \text{Sozialstaatsmoral} + \Gamma_3 * X_3 + \varepsilon_3$$

Signifikant negative Koeffizienten von Steuer- und Sozialstaatsmoral würden gesetzwidriges Verhalten bedeuten. In beiden Fällen wird ceteris paribus ein höherer Nettolohn erwartet.

Die endogenen Steuer- und Sozialstaatsmoral Variablen haben Halla und Schneider in Instrumental-Variablen umgeformt, indem sie beide auf die unabhängigen Variablen, wie viel Vertrauen die jeweils befragte Person dem Rechtssystem entgegenbringt, wie viel ehrenamtliche Freiwilligenarbeit sie leistet und wie viele Kinder die Befragten haben, regressieren. In Tabelle 3 sind die Ergebnisse der 2SLS-Schätzung für das Einkommen dargestellt: Es besteht kein signifikanter Einfluss der Steuermoral auf das Haushaltseinkommen, sehr wohl aber ein statistisch signifikanter negativer Einfluss der Sozialstaatsmoral. Eine um einen Punkt höhere Sozialstaatsmoral reduziert das Haushaltseinkommen um durchschnittlich 292 € pro Monat. Darüber hinaus haben Haushalte mit einem weiblichen Hauptverdiener ein um ca. 341 € niedrigeres Haushaltseinkommen pro Monat, wohingegen ein höherer sozioökonomischer Status (von den Befragten selbst angegeben, gemessen auf einer 4-Punkte-Skala) mit einem um 438 € höheren Einkommen pro Monat verbunden ist.

Tabelle 3: 2SLS-Schätzung des Einkommens

Unabhängige Variablen	Abhängige Variable: Einkommen
Steuermoral	1,610 (1,293)
Sozialstaatsmoral	-2,916** (1,309)
Dienstgrad des Hauptverdieners	0,113 (0,096)
Hauptverdiener ist weiblich	-3,410*** (0,475)
Sozioökonomischer Status	4,382*** (0,294)

1999 (=1, Dummy-Variable)	-0,066 (0,597)
Tirol	-1,827*** (0,728)
Konstante	12,477 (6,280)
Beobachtungen	1.887
Basmann-Statistik ^a	0,622
Sargan-Statistik ^b	0,620

Es gelten die Anmerkungen aus Tabelle 2. Quelle: Halla und Schneider (2005, 2014).

4. Schlussbemerkungen

Aus den Ergebnissen von Kapitel 3 sowie aus der Studie von Torgler und Schneider (2007), die hier nicht dargestellt ist, lassen sich die folgenden fünf Schlussfolgerungen ziehen:

- (1) Um den Fiskalbetrug zu verringern kann der Staat die Kontrollintensität erhöhen. Diese wirkt aber nur, wenn es aus der subjektiven Sicht der Steuerzahler wahrscheinlicher ist, dass sie „erwischt“ werden. Nur dann nimmt ihre Neigung zu Sozial-Betrug, als auch der Schwarzarbeit nachzugehen bzw. Steuern zu hinterziehen, ab. Die Höhe des Strafmaßes hat in den meisten Studien hingegen keinen eindeutigen signifikanten Einfluss.
- (2) Eine Politik der Abschreckung und des Misstrauens gegenüber dem Steuerzahler ist aber nicht der einzige Weg, die Bereitschaft der Bürger zu steigern, ihre Abgaben zu leisten und den Sozialstaat nicht zu betrügen. Ein anderer Weg führt über eine Verbesserung der staatlichen Leistungen. Für hochwertige aber auch quantitativ ausreichende Bereitstellung von öffentlichen Gütern und Leistungen zahlen die Bürger bereitwilliger Steuern als für schlechte Dienstleistungen.
- (3) Auch ein freundlicher, kundenorientierter Umgang der Finanzbehörden mit dem Steuerpflichtigen trägt zur Steuerehrlichkeit bei.
- (4) Darüber hinaus gewährleistet ein hohes Niveau an lokaler Autonomie und/oder an direkter Demokratie eine stärkere Berücksichtigung der Wähler-Präferenzen und erhöht damit die Identifikation der Bürger mit den staatlichen Institutionen. Dies wiederum dämpft ihre Nei-

gung zu Schwarzarbeit und Steuerhinterziehung, wie für die Schweiz gezeigt werden kann¹⁶.

- (5) Nicht zuletzt heben auch Nationalstolz, das Vertrauen gegenüber dem Rechtssystem und eine funktionierende Demokratie die Steuermoral.

Aus den Erkenntnissen von Kapitel 2 lassen sich die folgenden zwei Schlussfolgerungen ziehen: In vielen Ländern ist eine stark ausgeprägte Schattenwirtschaft mit viel Korruption verbunden. Eine große Schattenwirtschaft bedeutet, dass der Fiskus geringere Steuereinnahmen verbuchen kann, als es der Wirtschaftsleistung im weiteren Sinne entspricht. Dies wiederum verringert die Quantität und/oder Qualität der vom Staat bereitgestellten öffentlichen Gütern und Dienstleistungen sowie der Infrastruktur – mit der Nebenwirkung eines verminderten Anreizes für die Bürger, in der offiziellen Wirtschaft zu arbeiten. Wenn dann noch ein schwaches Rechtssystem und instabile wirtschaftliche Rahmenbedingungen dazu kommen, fördert dies die Korruption. Der hierfür zu betreibende Aufwand mit mehr Kontrollen und mehr staatlichem Aufwand, der wieder über Steuern finanziert wird, wirkt wie eine zusätzliche Abgabenbelastung, was dann zu noch mehr Schwarzarbeit führt und es entsteht ein Teufelskreis aus Schwarzarbeit und Korruption, wie dies in einigen Entwicklungsländern (beispielsweise Simbabwe oder Venezuela) zu beobachten ist.

Was kann getan werden?

- (1) Um die Korruption zu bekämpfen, sind in erster Linie die staatlichen Institutionen zu stärken, eine unabhängige Justiz zu schaffen und Pressefreiheit zu garantieren.
- (2) Weiterhin kann die Schattenwirtschaft verringert werden, indem die „schwarz“ erwirtschaftete Wirtschaftsleistung in die offizielle Wirtschaft überführt wird. Dies wird durch eine anreizorientierte Wirtschaftspolitik erreicht, mit der die Steuerlast der Bürger verringert und ihnen ganz allgemein mehr wirtschaftliche Freiheit gelassen wird.

16 Vgl. hierzu Torgler et al. (2010).

Referenzen

- Allingham, M.G. & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3–4), 323–338.
- Buehn, A. & Schneider, F. (2012). Corruption and the shadow economy: Like oil and vinegar, like water and fire?. *International Tax and Public Finance*, 19(1), 172–194.
- Choi, J. & Thum, M. (2005). Corruption and the shadow economy. *International Economic Review*, 46(3), 817–836.
- Dreher, A. & Schneider, F. (2009). Corruption and the shadow economy: An empirical analysis. *Public Choice*, 144(2), 215–277.
- EVS (2015). *European Values Study Longitudinal Data File 1981–2008 (EVS 1981–2008)* [Data set]. GESIS Data Archive. <https://doi.org/10.4232/1.12253>
- Feld, L.P., & Frey, B.S. (2004). Illegal, Immoral, Fattening or What?: How Deterrence and Responsive Regulation Shape Tax Morale. In *Marburg Working Papers on Economics* (No. 200426; Marburg Working Papers on Economics). Philipps-Universität Marburg, Faculty of Business Administration and Economics, Department of Economics (Volkswirtschaftliche Abteilung). <https://ideas.repec.org/p/mar/volksw/200426.html>.
- Feld, L.P. & Frey, B.S. (2007). Tax compliance as the result of a psychological tax contract: The role of incentives and responsive regulation. *Law and Policy*, 29(1), 102–120.
- Feld, L.P. & Schneider, F. (2010). Survey on the shadow economy and undeclared earnings in OECD countries. *German Economic Review*, 11(2), 109–149.
- Friedman, E., Johnson, S., Kaufmann, D. & Zoido-Lobaton, P. (2000). Dodging the grabbing hand: The determinants of unofficial activity in 69 countries. *Journal of Public Economics*, 76(4), 459–493.
- Halla, M., & Schneider, F. (2005). *Taxes and Benefits: Two Distinct Options to Cheat on the State?* (Working Paper No. 0505). Institut für Volkswirtschaftslehre, Johannes Kepler Universität Linz. <https://www.econstor.eu/handle/10419/73282>.
- Halla, M. & Schneider, F. (2014). Taxes and benefits: Two options to cheat on the state. *Oxford Bulletin of Economics and Statistics*, 76(3), 411–431.
- Medina, L. & Schneider, F. (2021). The evolution of shadow economies through the 21st century. In: C. Delechât & L. Medina (Hrsg.), *The Global Informal Workforce: Priorities for Inclusive Growth* (S. 10–69), Washington DC: International Monetary Fund.
- Meon, P. & Weill, L. (2010). Is corruption an efficient grease?. *World Development*, 38, 244–259.
- Poniatowski, G., Bonch-Osmolovskiy, M., Durán, J. M., Esteller-Moré, A., & Śmietanka, A. (2019). *Study and reports on the VAT gap in the EU-28 member states: 2019 final report* (CASE Reports No. 500). Center for Social and Economic Research (CASE). <http://hdl.handle.net/10419/227645>.
- Rose-Ackerman, S. (1997). *Corruption and development*. Annual Bank Conference on Development Economics, The World Bank, Washington DC.

- Schneider, F. (2015). Size and development of the shadow economy of 31 European and 5 other OECD countries from 2003 to 2014: Different developments?. *Journal of Self-Governance and Management Economics*, 3(4), 7–29.
- Schneider, F. (2021). Mehr Licht im Schatten? Neuere Ergebnisse über die Schattenwirtschaft. *Perspektiven der Wirtschaftspolitik*, 22(3), 225–246. <https://doi.org/10.1515/pwp-2021-0037>.
- Schneider, F. & Buehn, A. (2016). Shadow economy: Estimation methods, problems, results and open questions. *Open Economics*, 1(1), 1–29.
- Schneider, F. & Enste, D. (2000). Shadow economies: Size, causes and consequences. *Journal of Economic Literature*, 38(1), 73–110.
- Schneider, F., & Boockmann, B. (2021). *Die Größe der Schattenwirtschaft – Methodik und Berechnungen für das Jahr 2021* [Hintergrundpapier]. Institut für angewandte Wirtschaftspolitik (IAW), Universität Tübingen.
- Torgler, B. & Schneider, F. (2007). What shapes attitudes toward paying taxes? Evidence from multicultural European countries. *Social Science Quarterly*, 88, 443–470.
- Torgler, B. & Schneider, F. (2009). The impact of tax morale and institutional quality on the shadow economy. *Journal of Economic Psychology*, 30(2), 228–245.
- Torgler, B., Schneider, F. & Schaltegger, C.A. (2010). Local autonomy, tax morale, and the shadow economy. *Public Choice*, 144(1–2), 293–321.
- Williams, C. C., & Schneider, F. (2016). *Measuring the global shadow economy: The prevalence of informal work and labour*. Cheltenham & Northampton, MA: Edward Elgar Publishing.

Wettbewerbspolitik und Demokratie: Das Spannungsfeld von Konsumenten-, Produzenten- und Bürgerinteressen

Viktor J. Vanberg¹

Abstract

Politikempfehlungen, die eine angewandte Ökonomik in einem demokratischen Gemeinwesen machen kann, sollten mit Argumenten begründbar sein, die zeigen, dass die empfohlenen Maßnahmen den gemeinsamen Interessen der Bürgerschaft dienen. Was den Ordnungsrahmen der Wirtschaft anbelangt, so hat die Theorie der Wirtschaftspolitik seit Adam Smith ihre Empfehlung einer am Leitbild der Konsumentensouveränität orientierten Wettbewerbsordnung auf die These gestützt, dass die beteiligten Akteure zwar stets gleichzeitig als Konsumenten und Produzenten von den Spielregeln des Wirtschaftsgeschehens betroffen sind, dass es aber in ihrem *gemeinsamen* Interesse liegt, bei der Regelwahl ihren *Konsumenteninteressen* Priorität gegenüber ihren *Produzenteninteressen* einzuräumen. Der vorliegende Beitrag erörtert die diesem Leitbild zugrundeliegenden Argumente und zeigt auf, welche Komplikationen sich ergeben, wenn man berücksichtigt, dass Menschen nicht nur als Konsumenten und Produzenten, sondern auch als *Bürger* spezifische Interessen haben, und dass der in der Wirtschaftstheorie üblicherweise thematisierte Konflikt zwischen Produzenten- und Konsumenteninteressen von grundsätzlich anderer Natur ist als der Konflikt zwischen Bürgerinteressen und Konsumenteninteressen.

Keywords:

Wettbewerbspolitik, Demokratie, Konsumenteninteressen, Produzenteninteressen, Bürgerinteressen

1 Walter Eucken Institut, Freiburg; vanberg@eucken.de.

1. Einleitung

Praktische Wettbewerbspolitik, so wie jegliche menschliche *Praxis*, hat mit der Lösung von Entscheidungsproblemen zu tun, ist mit der Notwendigkeit konfrontiert, unter den in einer gegebenen Situation als verfügbar wahrgenommenen Handlungsalternativen eine Wahl zu treffen. Solche Wahlentscheidungen beruhen auf zwei Grundlagen, zum einen faktischen Annahmen darüber, welche Konsequenzen von den in Erwägung gezogenen Handlungsoptionen zu erwarten sind, und zum anderen einem normativen Maßstab, der es erlaubt, die jeweils zu erwartenden Konsequenzen nach ihrer Vorzugswürdigkeit zu beurteilen.²

Die Theorie der Wettbewerbspolitik als eine auf politische Anwendung ausgerichtete Wettbewerbsökonomik ist, wie jede Erfahrungswissenschaft, in ihrer Kompetenz auf überprüfbare Urteile zu den Fragen beschränkt, die die erste der oben genannten Grundlagen praktischer Wettbewerbspolitik betreffen. Ihre Autorität als Wissenschaft beansprucht sie für die „wenn-dann“-Hypothesen, die sie aufstellt, sie kann sie nicht in Anspruch nehmen, um Werturteile zu fällen, wie sie die praktische Politik als ihre zweite Grundlage erfordert. Nun ist der Möglichkeitsraum für Aussagen über potenzielle wettbewerbspolitische Maßnahmen und ihre Konsequenzen, auf die sich das Forschungsinteresse richten könnte, unbegrenzt, was die *Forschungspraxis* vor ein Entscheidungsproblem stellt, die Entscheidung darüber, auf welche Fragen man die – im Unterschied zu den möglichen Fragen *begrenzten* – Forschungskapazitäten konzentrieren sollte. Für eine Theorie der Wettbewerbspolitik, die nicht nur akademische Glasperlenspiele betreiben möchte, liegt es nahe, die Relevanz für die praktische Wirtschaftspolitik zum Kriterium für die Fokussierung ihres Forschungskriteriums zu wählen.

Von Relevanz für die praktische Wettbewerbspolitik kann eine Theorie der Wettbewerbspolitik aber nur dann sein, wenn sie über faktische Wirkungszusammenhänge informiert, die für die Herbeiführung von Ergebnissen genutzt werden können, von denen man annehmen kann, dass sie für die praktische Politik in der Tat *erstrebenswert* sind. Das bedeutet, dass eine um praktische Relevanz bemühte Theorie der Wettbewerbspoli-

2 Hoppmann (1967, S. 78f.): „Wettbewerbspolitik ist Wirtschaftspolitik zur Gestaltung von Marktprozessen. ... Von den zahlreichen möglichen Ausprägungen der Marktprozesse sind einige erwünscht, andere sind es nicht. ... Mit Hilfe der Wettbewerbsdefinition sollen bestimmte gesetzliche Vorschriften, die sogenannten Spielregeln, geschaffen werden, um die Marktprozesse in die gewünschten Bahnen zu lenken.“

tik mit Annahmen über die *Wünschbarkeit* der Konsequenzen politischer Maßnahmen, also mit *Werturteilen*, operieren muss. Dies bedeutet aber nicht, dass sie damit in Widerspruch zu ihrem Anspruch als wertfreie Erfahrungswissenschaft gerät. Für die betreffenden Werturteile nimmt sie nicht ihre wissenschaftliche Fachkompetenz in Anspruch. Sie übernimmt sie *hypothetisch*, als *hypothetische Werturteile*, in dem Sinne, dass sie darüber informiert, welche Maßnahmen für eine Politik geeignet sind, die sich an dem unterstellten Wünschbarkeitskriterium ausrichtet. Mit anderen Worten, sie stellt Behauptungen der folgenden Form auf: Wenn man Wirkungen des Typs Y erzielen möchte, dann sind Maßnahmen des Typs X zweckdienlich. Solche *bedingten (hypothetischen)* Handlungsempfehlungen sind keine *Werturteile* im Sinne *unbedingter (kategorischer)* Sollens-Aussagen. Es handelt sich vielmehr um reine Tatsachenbehauptungen die wie gewöhnliche empirische Hypothesen über Ursache-Wirkungs-Zusammenhänge überprüfbar sind. Sie sind falsifiziert, wenn Maßnahmen vom Typ X sich als nicht geeignet erweisen, Wirkungen vom Typ Y zu erzielen, sie sind unvollständig, wenn auch Maßnahmen anderen Typs solche Wirkungen erreichen können, und sie sind belanglos, wenn der Adressat, an den sie gerichtet sind, überhaupt nicht den Wunsch hat, die fraglichen Wirkungen realisiert zu sehen.

2. Wettbewerbspolitische Leitbilder und Demokratie

Wenn in Lehrbüchern zur Wettbewerbsökonomie und -politik von „Wettbewerbspolitischen Leitbildern“ die Rede ist,³ so geht es dabei um die Art von hypothetischen Werturteilen, deren Bedeutung für die Fokussierung des Forschungsinteresses im Vorhergehenden erörtert worden ist. Die postulierten „Leitbilder“ sollen einen normativen Referenzmaßstab abgeben, der es erlaubt, Aussagen darüber zu treffen, welche Maßnahmen für eine daran orientierte Wettbewerbspolitik zweckdienlich sind. Mit anderen Worten, sie ermöglichen es, bedingte (hypothetische) Handlungsempfehlungen für eine Wettbewerbspolitik zu geben, die das entsprechende Leitbild als Ziel verfolgen möchte.

Nun richten sich Empfehlungen zwangsläufig, ob nun explizit oder implizit, an einen Adressaten, von dem man annehmen kann, dass er an Ratschlägen der erteilten Art interessiert ist. In einem demokratischen Gemeinwesen kann es dafür nur zwei Kandidaten geben. Dies sind zum

3 Dazu Knieps (2005, S. 67–77).

einen die *politischen Amtsträger*, die durch Wahl dazu beauftragt und autorisiert worden sind, für das Gemeinwesen Entscheidungen zu treffen. Und dies sind zum anderen die *Bürger*, die, als seine Mitglieder, das demokratische Gemeinwesen als *Bürgergenossenschaft* konstituieren,⁴ und deren Votum daher die Quelle bildet, aus der jegliche politische Amtsautorität ihre Legitimation schöpft.

In Beiträgen zur Theorie der Wettbewerbspolitik wird in der Regel stillschweigend davon ausgegangen, dass die mit entsprechender Gesetzgebungskompetenz ausgestatteten Amtsinhaber die natürlichen Adressaten wettbewerbspolitischer Empfehlungen sind. Nun liegt es in der Tat nahe, die amtierenden Politiker als die *direkten* Adressaten von Politikempfehlungen zu betrachten. In einem demokratischen Gemeinwesen als einer Bürgergenossenschaft sind jedoch die Mitglieder-Bürger die *letztendlichen* Adressaten, denen es Argumente zu bieten gilt, weshalb sie erwarten können, dass die empfohlenen Maßnahmen ihren Interessen dienen. Genauer gesagt, es sind Gründe zu benennen, warum die betreffenden Maßnahmen im *gemeinsamen Interesse* der Bürgerschaft liegen.

Der letztendliche Test dafür, ob demokratische Politik den gemeinsamen Bürgerinteressen dient, ist ihre allgemeine *Zustimmungsfähigkeit*. Eine solche Zustimmungsfähigkeit wird man zwar nicht für tagespolitische Einzelentscheidungen fordern können, aber doch für den *allgemeinen Regelrahmen*, in dem und gemäß dem diese Entscheidungen getroffen werden. Die Zustimmung der Bürger zu den grundlegenden Verfassungsregeln ist für ein demokratisches Gemeinwesen unentbehrlich, das dem Anspruch gerecht werden soll, ein, in den Worten von John Rawls (1979, S. 105), „Unternehmen der Zusammenarbeit zum gegenseitigen Vorteil“ zu sein. Alle Bürger müssen die begründete Erwartung haben können, dass der allgemeine Regelrahmen den politischen Prozess auf längere Sicht so steuert, dass das Gesamtergebnis ihnen zum Vorteil gereicht, auch wenn es immer wieder zu Entscheidungen kommen wird, die ihren Interessen zuwiderlaufen.⁵

Der Anspruch zum „gegenseitigen Vorteil“ aller Beteiligten zu wirken, gilt auch für eine Wettbewerbspolitik, der es um die Gestaltung des *allgemeinen Ordnungsrahmens* geht, in dem sich wirtschaftlicher Wettbewerb entfalten soll und der seinerseits auch die Eingriffsmöglichkeiten der Poli-

4 Zum Verständnis des demokratischen Staates als Bürgergenossenschaft siehe Vanberg (2003).

5 Zu der hier skizzierten Sichtweise der konstitutionellen politischen Ökonomie siehe eingehender Vanberg (2005).

tik an Regeln bindet. Auch eine solche Wettbewerbspolitik muss in einem demokratischen Gemeinwesen die von ihr propagierte *Wettbewerbsordnung* mit Argumenten vertreten können, die geeignet sind, die Bürger von der Vorteilhaftigkeit des vorgeschlagenen Regimes gegenüber in Betracht kommenden Alternativen zu überzeugen.

Nun wirft die oben getroffene Unterscheidung zwischen den amtierenden Politikern als den *direkten* und den Bürgern als *letztendlichen* Adressaten von Politikempfehlungen in einem demokratischen Gemeinwesen die Frage auf, ob bei beiden Adressaten die gleiche Interessenlage unterstellt werden kann. Anders gesagt, kann man annehmen, dass die Gründe die für eine Regelung als im allgemeinen Bürgerinteresse liegend angeführt werden können, für politische Amtsträger zu einem handlungsleitenden Motiv werden? Idealerweise sollte man sich wünschen, dass der die Spielregeln der Politik festlegende Ordnungsrahmen in einem demokratischen Gemeinwesen die handlungsleitenden Interessen der Politiker mit den gemeinsamen Interessen der Bürger in Einklang zu bringen vermag, dass Politiker erwarten können, dass es für ihre Erfolgsaussichten im politischen Wettbewerb förderlich ist, sich für das „allgemeine Bürgerwohl“ oder die „gemeinsamen Bürgerinteressen“ einzusetzen.⁶

Es ist offenkundig, dass die tatsächlich vorfindbaren demokratischen Verfassungen sich nur in mehr oder minder unvollkommener Weise als geeignet erweisen, einen derartigen Gleichklang zwischen Politikerinteressen und gemeinsamen Bürgerinteressen herzustellen. Die Anreizkonstellationen, denen Politiker in dem Wettbewerb, in dem sie sich behaupten müssen, gegenüberstehen, bringen sie leicht in einen Entscheidungskonflikt zwischen dem, was für ihre Wiederwahlchancen oder sonstigen persönlichen Anliegen förderlich ist, und dem, was der Einsatz für die Gemeininteressen der Bürgerschaft von ihnen fordern würde. Statt Maßnahmen zu unterstützen, die den gemeinsamen Bürgerinteressen dienen würden, kann es für ihre eigenen Ziele erfolgversprechender sein, sich in den Dienst partikularer Sonderinteressen zu stellen.⁷

Dieser Entscheidungskonflikt stellt für Politiker eine persönliche Herausforderung dar, er ist aber vor allem ein *Systemproblem*, das nur durch

6 James M. Buchanan (1992) hat dieses Ordnungsideal in einem Aufsatztitel mit der Frage umschrieben: „How can constitutions be designed so that politicians who seek to serve ‘public interest’ can survive?“.

7 Die in der Public Choice Literatur unter dem Stichwort „rent-seeking“ und in der Tradition der Freiburger Schule mit dem Begriff der Privilegienvergabe umschriebenen Probleme weisen auf die hier angesprochenen Defizite im Ordnungsrahmen demokratischer Politik hin. Siehe dazu etwa Vanberg (2014, S. 351ff.).

geeignete Reformen der „Spielregeln der Politik“ und nicht durch Appelle an die persönliche Moral der politischen Amtsträger verlässlich behoben werden kann. In dem Maße, in dem die „Gemeinwohlorientierten“ unter letzteren in Nachteil geraten gegenüber ihren für gemeinschaftsbezogene Appelle „tauben“ Konkurrenten, wird der Selektionsprozess des politischen Wettbewerbs dazu führen, dass sich die Verteilung in der „Politikerpopulation“ über die Zeit zugunsten der letzteren verschiebt.

Für die Beratungsangebote, die eine angewandte Wettbewerbsökonomik machen kann, bedeutet der oben erörterte Konflikt, dass man zwischen *Politikerberatung* und *Politikberatung* im folgenden Sinne unterscheiden muss. *Politikerberatung* bedient die Nachfrage von Politikern nach Empfehlungen, die ihnen helfen, ihre Entscheidungen in einer Weise auszurichten, die ihre eigenen Erfolgchancen im politischen Wettbewerb erhöht. *Politikberatung* zielt darauf ab, Informationen darüber bereitzustellen, von welchen Maßnahmen man eine Förderung der gemeinsamen Bürgerinteressen erwarten kann.⁸ In den verbleibenden Teilen dieses Beitrages wird von der Annahme ausgegangen, dass *Politikberatung* im erläuterten Sinne die Aufgabe ist, der sich eine angewandte Wettbewerbsökonomik in einem demokratischen Gemeinwesen widmen sollte. Das stellt eine solche Wettbewerbsökonomik vor die Anforderung, ihre politischen Handlungsempfehlungen mit Argumenten zu begründen, die geeignet sind, die Bürger davon zu überzeugen, dass es in ihrem gemeinsamen Interesse liegt, den Empfehlungen zu folgen.

3. Konsumentensouveränität als Ordnungsideal

Als Adam Smith mit seiner „Untersuchung über Wesen und Ursachen des Reichtums der Völker“ ([1776] 2005) der „merkantilistischen“ Wirtschaftspolitik seiner Zeit Argumente für eine liberale Freihandelspolitik entgegensetzte, umschrieb er das von ihm vertretene Wirtschaftsordnungs-Ideal mit den Worten:

„Konsum ist der einzige Sinn und Zweck aller Produktion; und das Interesse des Produzenten sollte nur insoweit berücksichtigt werden, als es für die Förderung des Konsumenteninteresses nötig sein mag. Die Maxime ist so selbstverständlich, daß es unsinnig wäre, sie beweisen zu wollen. Doch im Merkantilssystem wird so gut wie immer das Interesse

8 Zu einer eingehenden Analyse dieser Unterscheidung siehe Cassel (2001).

des Konsumenten dem des Produzenten geopfert; und es dürfte die Produktion, nicht den Konsum, als Endzweck und Ziel aller Erwerbstätigkeit und allen Handels ansehen“ (ebd., S. 645).

Für die von Smith so umschriebene Leitvorstellung für die Wirtschaftspolitik hat William H. Hutt (1943, S. 215) den Begriff der *Konsumentensouveränität* geprägt, der ausdrücken soll, „that ultimate power to determine the use of resources which are ‚scarce‘ ... shall be vested in the people.“ Als Maxime für die Wirtschaftspolitik drückt er, so Hutt (ebd.), die Wertung aus, dass „the goodness or success of productive effort can be judged only in the light of consumer preferences“. Und er fordert eine Politik der *Sicherung von Wettbewerb*, die Schaffung *institutioneller Rahmenbedingungen*, die die Freiheit wirtschaftlicher Betätigung schützen.⁹

Das Konzept der Konsumentensouveränität gibt einen Maßstab vor, an dem gegebene wirtschaftliche Rahmenbedingungen zu messen sind, nämlich daran, inwieweit sie geeignet sind, sicherzustellen, dass Konsumentenentscheidungen als die maßgeblichen „Regler“ im Wirtschaftsprozess fungieren.¹⁰ Derselbe Maßstab ist gemeint, wenn die Vertreter der Freiburger Schule von Leistungswettbewerb als dem Steuerungsideal einer Wettbewerbsordnung sprechen, die dafür sorgt, dass „die Konsumenten über Art und Umfang der Produktion (bestimmen), wobei die Unternehmer letztlich ... in ihrem Auftrag handeln“ (Eucken, [1952] 1990, S. 163) oder, wie Franz Böhm (1980, S. 90) es bildhaft umschrieben hat, dass „der Konsument als Organist unmittelbar am Manual der Wirtschaftsorgel sitzt.“¹¹ Darin, die diesem Steuerungsideal dienlichen Rahmenbedingungen herzustellen und durchzusetzen, sahen die Freiburger die Aufgabe von Ordnungspolitik.

9 Hutt (1943, S. 215): „The arrangements required to make the ideal of consumers' sovereignty realisable exclude any idea of an acquiescent state. Hence economic freedom as I understand it does not mean laissez-faire in the popular sense. It simply means ... 'institutional planning', and when I talk of 'institutional planning' I am thinking of the fashioning of a framework within which free cooperation is possible.“

10 Hutt ([1936] 1990, S. 261): „According to the nature of the economic institutions tolerated or created by the State, we may regard consumers' sovereignty as receiving complete or incomplete expression; and it is under competitive institutions that we find its full and untrammelled realization“.

11 In ähnlich bildhafter Sprache hat Wilhelm Röpke (1960, S. 31) über das Ordnungsideal der Konsumentensouveränität gesagt, es gehe darum, sicherzustellen, „that the only road to business success is through the narrow gate of better performance in service of the consumer.“

Adam Smith betrachtete die politische Ökonomie „als Teil der Kunst des Staatsmannes oder Gesetzgebers“ (Smith, [1776] 2005, S. 443) und seine Empfehlung, „das Interesse des Produzenten sollte nur insoweit berücksichtigt werden, als es für die Förderung des Konsumenteninteresses nötig sein mag“ (ebd., S. 645), kann man entsprechend als einen an den Gesetzgeber adressierten Ratschlag für die Gestaltung des wirtschaftlichen Ordnungsrahmens betrachten. Für ein demokratisches Gemeinwesen bedeutet dies, der letztendliche Adressat der Smith'schen Empfehlung sind die Bürger, die das Gemeinwesen bildenden Mitglieder. Und das bedeutet wiederum, diese Empfehlung muss mit Argumenten begründet werden, die geeignet sind, die Bürger davon zu überzeugen, dass es in ihrem gemeinsamen Interesse liegt, einer entsprechenden Ordnung der Wirtschaft zuzustimmen.

Wenn bei Adam Smith, und generell bei der Rede von Konsumenten-souveränität, Produzenten- und Konsumenteninteressen gegenübergestellt werden, könnte der Anschein erweckt werden, es gehe um einen Interessenkonflikt zwischen unterschiedlichen Teilen der Bürgerschaft. In Wirklichkeit geht es vielmehr um einen *intra-personellen* Interessenkonflikt, sind alle (erwerbstätigen) Bürger doch sowohl als Produzenten – sei es als Unternehmer, als Anteilseigner oder als abhängig Beschäftigte – wie auch als Konsumenten am Wirtschaftsleben beteiligt.¹² Es gilt aber, ihnen Argumente zu bieten, warum es für sie ratsam ist, bei der Wahl der Wirtschaftsordnung ihren Interessen als Konsumenten Priorität einzuräumen und ihre Interessen als Produzenten hinten zu stellen. Anders gesagt, es gilt Argumente dafür zu bieten, warum es im gemeinsamen Interesse aller Bürger ist, als Produzenten die Bürden zu akzeptieren, die ihnen eine Ordnung des „Leistungswettbewerbs“ auferlegt und dafür als Konsumenten die Vorteile genießen zu können, die ihnen eine solche Ordnung verspricht.

Nun lässt sich in der Tat ein überzeugendes Argument dafür machen, dass die Bürger sich in ihrer doppelten Eigenschaft als Produzenten und Konsumenten von einem Regime der Wettbewerbsfreiheit eine günstigere Bilanz von Vorteilen und Nachteilen versprechen können als dies in einem protektionistischen Regime der Fall wäre. Die Nachteile, die sie als Produzenten in einer wettbewerblichen Ordnung aufgrund von Handlungen

12 Hutt ([1936] 1990, S. 257f.): „In regarding the individual as a consumer, we do not see him in his full relationship to society. He is also a producer. But as producer he is the servant of the community. ... As a 'consumer' each directs. As a 'producer' each obeys.”

ihrer Konkurrenten zu gewärtigen haben, werden durch die „Generalkompensation“ (von Weizsäcker, 1998, S. 280),¹³ die sie in einer solchen Ordnung als Konsumenten erfahren, mehr als ausgeglichen. Die Überzeugungskraft dieses Arguments ist allerdings daran gebunden, dass man die Vorteils-Nachteils-Bilanz in einer allgemein wettbewerblichen Ordnung mit derjenigen eins durchgängig protektionistischen Regimes vergleicht. In der demokratischen Alltagspolitik stehen Bürger aber nicht vor einer Wahl zwischen diesen beiden Alternativen. Vielmehr geht es typischerweise um die Frage, ob man sich als Produzent um Wettbewerbsschutz (oder Kompensation für Wettbewerbsfolgen) zugunsten des eigenen Gewerbezweiges einsetzt oder auf solche Lobbytätigkeit verzichtet. Und in dieser Frage sehen Bürger sich dem in der folgenden Matrix dargestellten Gefangenendilemma ausgesetzt.

Abbildung 1: Protektionismus-Dilemma

		die anderen Produzenten	
		verzichten auf Lobbying	suchen Protektion
Produzent X	verzichtet auf Lobbying	Z	S
	sucht Protektion	B	D

Aus Sicht von Produzent X: $B > Z > D > S$

Quelle: Eigene Darstellung

Zwar ist aus der Sicht des Einzelnen Z (alle verzichten auf Lobbying) gegenüber D (alle suchen Protektion) vorzuziehen, aber die attraktivste Situation wäre es, als einziger ein Produktionsprivileg zu genießen (B) und am ungünstigsten wäre es, sich als einziger „Dummer“ unter ansonsten protektionsbegünstigten Produzenten zu finden. Dieses Dilemma, das aus der Sicht des Einzelnen die Alternative, sich um den Wettbewerbsschutz zu bemühen, in jedem Fall als die attraktivere erscheinen lässt, ist, wie auch von Weizsäcker (1998, S. 279f.) deutlich macht, nur durch eine vor-

13 Das Argument gilt in gleicher Weise für Protektionismus in der Form von Schutz gegen Wettbewerb wie in der Form der Kompensation für Wettbewerbsfolgen. Im Einzelfall könnte zwar, so stellt von Weizsäcker (1998, S. 280) fest, der Schutz vor unerwünschten Wettbewerbswirkungen „dem einzelnen Bürger helfen. In der Summe und unter Berücksichtigung des gleichzeitigen Verzichts der anderen Bürger ... profitiert er von diesem gemeinsamen Verzicht.“

gängige Verfassungsentscheidung zu lösen, die Regierung und Gesetzgeber auf Diskriminierungsfreiheit verpflichtet und ihnen die Vergabe von Produktionsprivilegien untersagt.¹⁴ Ebendies zu erreichen war das Ziel, das die Vertreter der Freiburger Schule mit ihrem Einsatz für eine auf *Ordnungspolitik* verpflichtende *Wirtschaftsverfassung* anstrebten,¹⁵ und dieses Ziel haben Vertreter der konstitutionellen politischen Ökonomie im Auge, wenn sie von „politics by principle“ (Buchanan & Congleton, [1998] 2003) sprechen.

4. Konsumentensouveränität und Bürgersouveränität

In seiner *Untersuchung über Wesen und Ursachen des Reichtums der Völker* betont Adam Smith insbesondere den Zusammenhang zwischen Wohlstand und der Größe des Marktes. Die wesentliche und nachhaltige Quelle für Wohlstandswachstum liegt nach Smith in der Steigerung der Produktivität menschlicher Arbeit. Und zu letzterer stellt er fest:

„Die bedeutendste Steigerung der Produktivität der Arbeit und der Großteil der Geschicklichkeit, Fertigkeit und Umsicht, mit der sie überall eingesetzt oder verrichtet wird, dürfte die Wirkung der Arbeitsteilung gewesen sein“ (Smith, [1776] 2005, S. 89).

Und der Grad der Arbeitsteilung, so argumentiert er, ist wiederum abhängig von der Ausdehnung des Marktes:

„Da die Möglichkeit zu tauschen Anlaß zur Arbeitsteilung gibt, muß das Ausmaß dieser Teilung stets durch das Ausmaß jener Möglichkeit begrenzt sein, mit anderen Worten: durch die Größe des Marktes. Wenn der Markt sehr klein ist, kann niemand einen Anreiz haben, sich gänzlich einer einzigen Beschäftigung zu widmen, weil es an der Möglichkeit fehlt, den ganzen, über seinen eigenen Verbrauch hinausgehenden, überschüssigen Teil des Ertrages seiner eigenen Arbeit nach Bedarf gegen entsprechende Teile des Arbeitsertrages anderer Menschen einzutauschen“ (Smith, [1776] 2005, S. 101).

14 Wie M. Olson ([1965] 1968) eingehend dargelegt hat, stehen Einzelne als Mitglieder einer Gruppe von Produzenten (wie generell Mitglieder einer Interessengruppe) bei der Entscheidung, sich für die Belange der Gruppe einzusetzen, ebenfalls vor einem Gefangenendilemma. Bei der Lösung solcher Dilemmata ist die Organisationsfähigkeit der jeweiligen Gruppe von entscheidender Bedeutung.

15 Siehe dazu Vanberg (1997a).

Die Folgerung aus dieser für die Begründung des Freihandels grundlegenden Argumentationskette ist, dass Menschen sich von einer Vergrößerung des Marktes – folgt man der obigen Argumentation zur Abwägung von Produzenten- und Konsumenteninteressen – eine Verbesserung ihrer Lage versprechen können. Und dies wiederum sollte vermuten lassen, dass sie die mit dem Begriff der *Globalisierung* umschriebene weltweite Vernetzung von Produktion und Handel willkommen heißen.

Nun hat jedoch die durch eine deutliche Beschleunigung solcher Vernetzung gekennzeichnete Entwicklung seit dem Zusammenbruch des Sowjetimperiums gezeigt, dass die mit der Globalisierung einhergehenden verbesserten Konsummöglichkeiten zwar durchaus geschätzt werden, dass aber gerade in den entwickelten Volkswirtschaften auch vermehrt globalisierungskritische Stimmen laut werden und im Parteienwettbewerb ihren Widerhall finden. Das darin zum Ausdruck kommende Unbehagen an der Globalisierung hat sicherlich nicht zuletzt mit verstärktem Wettbewerbsdruck zu tun, dem Menschen als Produzenten in einem weltweiten Markt ausgesetzt sind, zumal dieser erhöhte Wettbewerbsdruck sich asymmetrisch auswirkt. Er begünstigt die Anbieter von Kapital und qualifizierter Arbeit, die weltweit vermehrt nachgefragt werden, und er wirkt sich besonders ungünstig für die Anbieter gering oder unqualifizierter Arbeit in den entwickelten Volkswirtschaften aus, die mit einem weltweit deutlich erhöhten Angebot konkurrieren müssen. Was allerdings die aus dem globalisierungsbedingten erhöhten Wettbewerbsdruck resultierenden protektionistischen Interessen anbelangt, so bedeuten sie für das Problem der Abwägung zwischen ihren Interessen als Produzenten und Konsumenten, die die Bürger eines demokratischen Gemeinwesens bei der Entscheidung über ihre Wirtschaftsordnung treffen müssen, keine grundlegende Änderung. Für die traditionellen protektionistischen Interessen gilt weiterhin die oben dargelegte Argumentation der Generalkompensation, stehen dem erhöhten Wettbewerbsdruck, dem Bürger sich als Produzenten ausgesetzt finden, doch erhöhte Vorteile gegenüber, die sie als Konsumenten genießen können, sodass man weiterhin eine positive Nettobilanz unterstellen kann. Die Befriedigung dieser Protektionsinteressen ist nur solange von Vorteil, wie sie als Privileg an wenige gewährt wird. Was die asymmetrisch betroffenen Anbieter gering qualifizierter Arbeit anbelangt, so könnte eine Bürgerschaft diesem Problem mit einem begleitenden *Sozialpakt* begegnen, der Globalisierungsgewinne nutzt, um den Betroffenen durch Qualifizierungs- und Weiterbildungsprogramme sowie andere Unterstützungsmaßnahmen beizustehen.

So bedeutsam die oben erörterten, von Bürgern in ihrer Eigenschaft als *Produzenten* vertretenen, protektionistischen Interessen auch fraglos für das

Verständnis globalisierungskritischer Tendenzen sind, sie sind doch nicht die einzige Quelle, aus der letztere sich speisen. Wenn Walter Eucken ([1940] 1989, S. 240) von einer „funktionsfähigen und menschenwürdigen Ordnung“ sagt, dass in ihr „die Knappheit an Gütern ... so weitgehend wie möglich und andauernd überwunden ... (und) zugleich ... ein selbstverantwortliches Leben möglich“ sein solle, dann deutet er damit an, dass der Maßstab, nach dem Menschen die Güte der Ordnung beurteilen, in der sie zu leben wünschen, sich nicht auf ihre Interessen als Produzenten und Konsumenten beschränkt.

In der Tat kann man einen Teil des Unbehagens an der Globalisierung nur verstehen, wenn man in Rechnung stellt, dass in ihm nicht allein protektionistische Privilegieninteressen zum Ausdruck kommen. Offenkundig haben Bürger als Mitglieder eines Gemeinwesens Interessen an der Bewahrung bestimmter Qualitäten ihrer Lebensumwelt, die sie durch Freizügigkeit der grenzüberschreitenden Bewegung von Gütern, Kapital und Personen gefährdet sehen. Insbesondere zuwanderungskritische Einstellungen spielen bei den Vorbehalten gegen die Globalisierung eine offenkundige Rolle, eine Rolle, die von gewissen Parteien zur Wählerwerbung genutzt wird. Es ist bedenkenswert, das F.A. Hayek, dem Sympathien für solche Parteien gänzlich ferngelegen hätten, sich recht nachdenklich zum Migrationsproblem äußert, wenn er feststellt:

„Freedom of migration is one of the widely accepted and wholly admirable principles of liberalism. But should this generally give the stranger a right to settle down in a community in which he is not welcome? ... Or ought it to be an offence if they voluntarily agree not to do so? Swiss and Tyroles villages have a way of keeping out strangers which neither infringe nor rely on any law. Is this anti-liberal or morally justified? For established old communities I have no certain answers to these questions” (Hayek, 1979, S. 195).

Gewiss wird man solchen Bedenken Argumente entgegenhalten können, die – ähnlich wie bei der Abwägung von Produzenten- und Konsumenteninteressen – darauf hinweisen, dass den mit den globalisierungsbedingten Umweltveränderungen verbundenen Einbußen Vorteile gegenüberstehen, die durchaus zu einer positiven Gesamtbilanz führen können. Allerdings liegt der entscheidende Unterschied zwischen den beiden Konstellationen darin, dass es im ersten Fall darum geht, dass Bürger ihre Privilegieninteressen als Produzenten gegen ihre konsensfähigen Interessen als Konsumenten abzuwägen haben, während es im zweiten Fall um einen Konflikt zwischen ihren (globalisierungsfreundlichen) *Konsumenteninteressen* und ihren (globalisierungskritischen) *Bürgerinteressen* i. S. ihrer Interessen an

der Bewahrung geschätzter Qualitäten ihrer Lebensumwelt geht. Solche protektionistischen Bürgerinteressen sind von grundlegend anderer Natur als protektionistische Produzenteninteressen und man kann ihnen nicht mit denselben Argumenten begegnen, die man letzteren entgegenhalten kann.

Der Föderalismusexperte John Kincaid hat die hier in Frage stehende Interessenkonstellation als einen Konflikt zwischen „*consumership*“ und „*citizenship*“ umschrieben. Damit bezeichnet er den Konflikt zwischen unseren Interessen als *Konsumenten* an der Wahlfreiheit, die ein großer Markt zu bieten vermag, und den Interessen, die wir als *Bürger* daran haben, als bedeutsam erachtete Charakteristika unseres Gemeinwesens nach unseren Wünschen gestalten zu können.¹⁶ Als Konsumenten nehmen wir unsere *Privatautonomie* wahr, also die Handlungsfreiheiten, die wir als *Privatrechts-subjekte* genießen. Als Bürger können wir die *Mitbestimmungsrechte* ausüben, über die wir als Mitglieder eines demokratischen Gemeinwesens verfügen. In Kincaid's (1994, S. 37) Worten:

„Consumership refers to the empowerment derived from access to the global marketplace for goods and services. Citizenship refers to the empowerment derived from participation in a self-governing political community having a distinct identity. ... The problem is that the requisites of consumership can diminish the citizenship opportunities provided by constituent political jurisdictions, while the requisites of citizenship can diminish the consumership opportunities provided by a common market.”

Die meisten Menschen, so stellt Kincaid (1992, S. 34) fest, wünschen sich beides: „[they] wish to be citizens of an identifiable place, ... but they also desire the means and freedom to consume the goods and services available on the world market“ (ebd., S. 31). Die sich im Unbehagen an der Globalisierung äußernde Spannung zwischen diesen beiden Wünschen ist, wie Kincaid deutlich macht, ein Problem, das sich generell in Föderationen (wie etwa der Europäischen Union) bei der Frage der Kompetenzverteilung zwischen der föderalen Ebene und den einzelstaatlichen bzw. lokalen Einheiten stellt:

„This tension arises from the dynamics of consumerism, which increases incentives for citizens to demolish jurisdictional barriers that inhibit the movement of goods and services across the boundaries of the con-

16 Zu einer eingehenderen Erörterung dieses Konflikts siehe Vanberg (1997b).

stituent governments within the federation's common market. Free trade requires diminution of the powers of self-government exercised by the constituent governments because variable exercises of such powers ... would fragment the marketplace and frustrate consumers. Yet these powers of regional and local self-government are prized by locals who seek to maintain communal values and public policies that suit their preferences as local rather than national citizens. The more that citizens pursue consumership interests in the local marketplace by levelling jurisdictional barriers, however, the more they reduce the effective scope of their local citizenship rights. Thus, while local citizenship has centrifugal effects on power in a federal system, national consumership has centripetal effects" (Kincaid, 1995, S. 261).¹⁷

5. Schluss

Seit Adam Smith sein Argument zum Zusammenhang von Arbeitsproduktivität, Arbeitsteilung, Größe des Marktes und Schaffung von Wohlstand dargelegt hat, haben Ökonomen ihre Aufmerksamkeit bei ihren Überlegungen zur Gestaltung wettbewerblicher Ordnung auf den Konflikt zwischen Produzenten- und Konsumenteninteressen konzentriert. Und sie haben sich in der Regel der von Adam Smith gezogenen Folgerung angeschlossen, dass eine Bürgerschaft gut beraten ist, ihren Konsumenteninteressen Vorrang vor ihren protektionistischen Produzenteninteressen einzuräumen, da man ja schließlich nur produziert, um zu konsumieren. In diesem Sinne haben sie denn auch die in der Politik verbreitete Neigung, die Nachfrage nach Wettbewerbsschutz zu bedienen, typischerweise als Indikator für einen Defekt der politischen Ordnung gewertet, der Unfähigkeit, die Gemeininteressen der Bürger gegen die Privilegieninteressen von Produzentengruppen durchzusetzen.

Nun geben die globalisierungskritischen Strömungen, die über die letzten Jahrzehnte in vielen Ländern an politischer Dynamik gewonnen und zu Verwerfungen in der Parteienlandschaft geführt haben, Anlass, darüber nachzudenken, ob die Konzentration auf den von Produzentenin-

17 Kincaid (1995, S. 260): „A federal polity ... faces a ... balancing challenge ... between individual liberty and communitarian liberty, or, in shorthand, the rights of persons (personal freedoms) versus the rights of places (states' rights). Federalism's protections of the rights of the places do not necessarily protect the rights of persons, while protections of the rights of persons do not necessarily protect the rights of places.“

teressen getriebenen Protektionismus nicht von Problemen ablenkt, die in den traditionellen Argumenten zugunsten einer Freihandelsordnung keine Beachtung finden. Bei der Entscheidung über die Gestaltung der wirtschaftlichen Ordnung ihres Gemeinwesens müssen Bürger nicht nur ihre Interessen als Produzenten und Konsumenten gegeneinander abwägen. Sie stehen auch vor der Notwendigkeit der Abwägung zwischen ihrem Interesse, als Konsumenten die Vorzüge des Zugangs zu einem großen Markt nutzen zu können, und ihrem Interesse daran, gemeinsam mit ihren Mitbürgern über Dinge entscheiden zu können, die für die Lebensqualität, die ihnen ihr Gemeinwesen bietet, von Bedeutung sind. Die Abwägungen, die in dem einen und dem anderen Fall zu treffen sind, haben eine unterschiedliche Qualität, und in einem demokratischen Gemeinwesen erfordert der Umgang mit ihnen andersartige Argumente.

Wenn sich etwa die Bürgerschaft einer Gemeinde gegen die Eröffnung eines McDonald's Restaurants auf dem Marktplatz ihres Ortes wendet, weil sie das historische Ortsbild erhalten will, so ist dies etwas anderes, als wenn die örtlichen Gastronomen die Niederlassung eines unliebsamen Konkurrenten verhindern wollen. Und die Abwägungsprobleme, vor die eine Bürgerschaft bei Entscheidungen über die *Art von Ordnung*, in der sie leben möchte, durch Probleme des einen und des anderen Typs gestellt ist, sind von grundlegend unterschiedlicher Natur. Im zweiten Fall geht es darum, ob eine Ordnung wünschenswert ist, in der von Fall zu Fall bestimmten Produzentengruppen Wettbewerbsschutz gewährt werden kann, oder ob es in ihrem gemeinsamen Interesse liegt, sich an eine generelle Regel zu binden, die jegliche Privilegienvergabe ausschließt. Beim ersten Problemtyp geht es für sie darum abzuwägen, ob es für sie insgesamt vorteilhafter ist, sich in einer Weise in einem umfassenderen Markt einzubinden, die es ihnen unmöglich machen würde, ihren im Beispiel genannten Wunsch nach Bewahrung des historischen Ortsbildes durchzusetzen, oder ob sie es vorziehen, sich für gewisse Sachverhalte die Kompetenz vorzubehalten, eigenständige Regelungen zu treffen.

Eine Wettbewerbsökonomik, die die Bürger demokratischer Gemeinwesen als ihre letztendlichen Adressaten betrachtet, wird bei ihren (bedingten) ordnungspolitischen Empfehlungen dem Unterschied zwischen den beiden Problemtypen Rechnung tragen müssen.

Referenzen

- Böhm, F. (1980). Wirtschaftsordnung und Staatsverfassung. In E.-J. Mestmäcker (Hrsg.), *Freiheit und Ordnung in der Marktwirtschaft* (S. 53–103) Baden-Baden: Nomos.
- Buchanan, J. M. (1992). How Can Constitutions Be Designed So That Politicians Who Seek To Serve 'Public Interest' Can Survive?. *Constitutional Political Economy*, 4, 1–6.
- Buchanan, J. M. & Congleton, R. D. ([1998] 2003). *Politics by Principle, Not Interest – Towards Nondiscriminatory Democracy*. Indianapolis: Liberty Fund.
- Cassel, S. (2001). *Politikberatung und Politikerberatung*. Bern et al.: Paul Haupt.
- Eucken, W. ([1940] 1989). *Die Grundlagen der Nationalökonomie*, 9. unveränderte Aufl. Berlin et al.: Springer.
- Eucken, W. ([1952] 1990). *Grundsätze der Wirtschaftspolitik*, Neuausgabe der 6., durchgesehenen Auflage. Tübingen: Mohr Siebeck.
- Hayek, Friedrich A. (1979). *The Political Order of a Free People*, London and Henley: Routledge & Kegan Paul.
- Hoppmann, E. (1967). Wettbewerb als Norm der Wettbewerbspolitik. *ORDO – Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft*, 18, 77–94.
- Hutt, W. H. (1943). *Plan for Reconstruction: A Project for Victory in War and Peace*. London: Kegan Paul.
- Hutt, W. H. ([1936] 1990). *Economists and the Public – A Study of Competition and Opinion*. New Brunswick & London: Transaction Publishers.
- Kincaid, J. (1992). Consumership versus citizenship: is there wiggle room for local regulation in the global economy?. In: B. Hocking (Hrsg.), *Foreign Relations and Federal States* (S. 27–47). London & New York: Leicester University Press.
- Kincaid, J. (1994). Economic Union and Federal Diversity: Should a Coca-Cola Bottle Carry 51 Health Warnings and, If So, in What Language?. In: A. Mullins & C. Saunders (Hrsg.), *Economic Union in Federal Systems* (S. 37–68). Sydney: The Federation Press.
- Kincaid, J. (1995). Liberty, Competition, and the Rise of Coercion in American Federalism. In L. Gerken (Hrsg.), *Competition Among Institutions* (S. 259–281). London: Macmillan.
- Knieps, G. (2005). *Wettbewerbsökonomie*, 2. Auflage. Berlin & Heidelberg: Springer.
- Olson, M. ([1965] 1968). *Die Logik des kollektiven Handelns: Kollektivgüter und die Theorie der Gruppen*. Tübingen: Mohr Siebeck.
- Rawls, J. (1979). *Eine Theorie der Gerechtigkeit*. Frankfurt a.M.: Suhrkamp.
- Röpke, W. (1960). *A Humane Economy – The Social Framework of the Free Market*. Chicago, Ill.: Henry Regnery.
- Smith, A. ([1776] 2005). *Untersuchung über Wesen und Ursachen des Reichtums der Völker*. Aus dem Englischen übersetzt von Monika Streissler, Herausgegeben und eingeleitet von Erich W. Streissler. Tübingen: Mohr Siebeck.

- Vanberg, V. J. (1997a). Die normativen Grundlagen von Ordnungspolitik. *ORDO – Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft*, 48, 707–726.
- Vanberg, V. J. (1997b). Subsidiarity, Responsive Government and Individual Liberty. In K. W. Nörr und T. Oppermann (Hrsg.). *Überlegungen zur Subsidiarität* (S. 253–269). Tübingen: Mohr Siebeck.
- Vanberg, V. J. (2003). Konsumentensouveränität und Bürgersouveränität: Steuerungsideale für Markt und Politik. *Politische Vierteljahresschrift*, Sonderheft 34 Politik und Markt, 48–65.
- Vanberg, V. J. (2005). Market and state: the perspective of constitutional political economy. *Journal of Institutional Economics*, 1, 23–49.
- Vanberg, V. J. (2014). Liberalismus und Demokratie: Zu einer vernachlässigten Seite der liberalen Denktradition. *ORDO – Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft*, 65, 345–374.
- Weizsäcker, Christian von (1998). Das Gerechtigkeitsproblem in der Sozialen Marktwirtschaft. *Zeitschrift für Wirtschaftspolitik*, 47, 257–288.

Die Anti-Stau-Maut

*Achim Wambach*¹

Abstract

2014 veröffentlichte der Wissenschaftliche Beirat beim Bundesministerium für Wirtschaft und Klimaschutz unter Federführung von Prof. Dr. Günter Knieps ein Gutachten zu „Engpassbasierte[r] Nutzerfinanzierung und Infrastrukturinvestitionen in Netzsektoren“. Darin empfahl der Beirat verkehrsabhängige Straßennutzungsgebühren, regionale Preise im Stromnetz, und qualitäts-differenzierende Preise der Internetnutzung. Der vorliegende Beitrag geht der Frage nach, was sieben Jahre später aus diesen Empfehlungen geworden ist.

Keywords:

Staumaut, Netzengpass, Verkehr, Elektrizität, Telekommunikation

1. Einleitung

Das Gutachten des Wissenschaftlichen Beirats beim Bundesministerium für Wirtschaft und Energie zur Bepreisung von Engpässen in den Sektoren Verkehr, Elektrizität und Telekommunikation und Internet wurde an einem Vormittag im Dezember 2014 wie üblich der Presse im Bundeswirtschaftsministerium vorgestellt. Die Empfehlungen des Gutachtens bezogen sich auf eine verkehrsabhängige Straßennutzungsgebühr, Preise im Stromnetz, die die Engpässe im Netz berücksichtigen, und auf differenzierte Preise für die Internetnutzung abhängig von der benötigten Qualität der Signaldurchleitung. Die Empfehlungen basierten auf der Erkenntnis, dass eine Infrastruktur, die an ihre Grenzen stößt, nicht wie ein öffentliches Gut behandelt werden sollte, da eine wesentliche Eigenschaft eines

1 Präsident des ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim; achim.wambach@zew.de
Ich bedanke mich bei Christian Mix und Dr. Ana Helena Palermo Kuss für die Unterstützung bei der Erstellung des Beitrags.

öffentlichen Gutes, die „Nicht-Rivalität“ im Konsum, nicht mehr gegeben ist. Die sogenannte Nicht-Rivalität ist dann gegeben, wenn der Konsum einer Person den Konsum anderer Personen nicht beeinträchtigt. Eine wenig befahrene Straße weist diese Eigenschaft auf, da ein zusätzliches Fahrzeug die anderen auf dieser Straße nicht beeinträchtigt. Kommt es allerdings zum Stau, ist diese Eigenschaft eines öffentlichen Gutes nicht mehr gegeben: In einer solchen Situation behindert jedes Fahrzeug die dahinter fahrenden Fahrzeuge.

Es mag dem sperrigen Titel geschuldet sein – *Engpassbasierte Nutzerfinanzierung und Infrastrukturinvestitionen in Netzsektoren* – aber auf jeden Fall war die Resonanz in der Pressekonferenz, die Prof. Günter Knieps als Federführender der Arbeitsgruppe zur Vorbereitung des Gutachtens und ich, in meiner damaligen Funktion als Vorsitzender des Beirats, durchführten, gering. Das sollte sich aber in den folgenden Stunden desselben Tages ändern.

Die Bild-Zeitung hatte zwischenzeitlich von dem Gutachten Wind bekommen. Unter der Schlagzeile „EXTRA-GEBÜHR IN STOSSZEITEN – Regierungs-Experten fordern Stau-Maut auf Autobahnen“ (Bild, 2014) machte sie Stimmung. Günter Knieps wurde zum gefragten Interviewpartner. In der Bild-Zeitung wurde er zitiert: „Unser Ziel ist es, die Zahl der Staus zu senken. [...] Wer unbedingt zur Stoßzeit unterwegs sein will, muss dann auch mehr bezahlen.“ Der damalige Verkehrsminister Alexander Dobrindt (CSU) trat vor die Kamera und bezog Stellung. Er sprach sich allerdings gegen eine solche Maut aus. Auch der damalige Wirtschaftsminister Sigmar Gabriel (SPD) äußerte sich: Er begrüßte zwar unseren Vorschlag, fürchtete aber, dass eine „stauabhängige Straßennutzungsgebühr“ Pendler erheblich zusätzlich belasten könnte.

Nach dem Vorstoß der Bild-Zeitung war das Medienecho gewaltig. Selbst die Tiroler Nachrichten berichteten von unserem Gutachten (s. Tiroler Tageszeitung, 2014). Auch Amüsantes war dabei. Ein Sender befragte Personen, die im Stau standen, was sie denn von dem Vorschlag der Wissenschaftler hielten, jetzt auch noch eine Gebühr zu zahlen. Die Begeisterung hielt sich verständlicherweise in Grenzen. Inhaltlich wurde das Gutachten allerdings auf den Vorschlag der Stau-Maut verkürzt; die Empfehlungen zu Strompreisen und zu differenzierten Preisen der Internetnutzung wurden nicht aufgegriffen.

Heute ist die Aufregung vergessen. Das Thema Stau-Maut (oder besser Anti-Stau-Maut), das in akademischen Kreisen immer wieder vorgebracht wurde, fand sich nach diesen Tagen nicht mehr in den Schlagzeilen wieder. Eine Suche nach „Staumaut“ bei Google Trends ergibt einen einsa-

men Peak im Jahr 2014, dem Jahr des Gutachtens. Das Thema ist aber, wie die beiden anderen auch, nach wie vor aktuell und drängend.

Dieser Artikel geht der Frage nach, was sich seit 2014 bei diesen Themen getan hat und wo wir heute stehen. In Anlehnung an das damalige Gutachten wird zunächst auf die ökonomischen Grundlagen in einem Abschnitt „Zur Ökonomie von Engpassmanagement und Netzausbauentscheidungen“ eingegangen. Im darauf folgenden Abschnitt werden für die jeweiligen Sektoren „Verkehr“, „Elektrizität“ und „Telekommunikation und Internet“ die damaligen Vorschläge aufgegriffen und die aktuelle Situation besprochen.

Es zeigt sich, dass im Verkehr der Bedarf nach einer „intelligenten Maut“ (mit diesem Ausdruck wurde Günter Knieps damals in der Presse zitiert) nach wie vor hoch ist und dass, auch technologiebedingt, umsetzungsfähige Vorschläge dazu vorliegen, die auch Interesse über die Wirtschaftswissenschaften hinaus finden. So wurde erst 2018 in einem Artikel in der Fachzeitschrift *Nature* darauf eingegangen (s. Cramton et al., 2018). Im Stromsektor wird das Problem der Netzengpässe immer virulenter, der Netzausbau stockt. Regionale Preise, die die Engpässe im Netz berücksichtigen, werden allorten diskutiert. Neben dem Vorschlag der EU Kommission zu mehreren Strompreiszonen für Deutschland ist auch die sogenannte generalisierte Merit Order, die von Günter Knieps (2013) entwickelt wurde, eine Alternative, die sowohl im Gutachten des Wissenschaftlichen Beirats als auch in der neulich veröffentlichten Studie aus der Schriftenreihe „Energiesysteme der Zukunft“ (s. Weyer & Müsgens, 2020) eingebracht wurde. Sie hätte den großen Vorteil, differenziert auf die immer wieder unterschiedlichen Engpässe im Stromnetz einzugehen.

Während hinsichtlich der Straßenmaut und den differenzierten Strompreisen wenig gemacht wurde, aber eine Bewegung zu konstatieren ist, ist die Lage bei den differenzierten Preisen für die Internetnutzung mittlerweile zumindest rechtlich verankert. Netzneutralität im Internet, also ein Verbot von differenzierten Preisen, wurde – entgegen den Vorschlägen des Gutachtens – festgeschrieben, allerdings mit Ausnahmen.

2. Zur Ökonomie von Engpassmanagement und Netzausbauentscheidungen

Straßen, das Stromnetz, das Telekommunikationsnetz sowie das Internet gehören zur Infrastruktur einer Gesellschaft. Diese Netze haben, solange sie nicht überlastet sind, die Eigenschaft der „Nicht-Rivalität“. Wenn das Netz allerdings an seine Grenzen stößt, ist diese Nicht-Rivalität nicht mehr

gegeben: Auf der Straße bildet sich Stau, und jede weitere Person, die in den Stau hineinfährt, verlängert den Stau für die dahinter kommenden Personen. Wenn die Anwohner einer ganzen Straße im Homeoffice versuchen, an Videokonferenzen teilzunehmen, geschieht ein analoges Phänomen: Die Datenrate bei jedem einzelnen Haushalt wird gedrosselt.

Das Vorliegen solcher Engpässe ist in einem Netz fast immer zwingend: Wenn das Netz so ausgebaut ist, dass nie ein Engpass auftritt, wurde das Netz vermutlich zu großzügig ausgelegt. Bei dem Vorliegen von Engpässen stellen sich zwei Fragenkomplexe:

- 1) Wie sollte man mit dem Engpass umgehen? Wer wird nach welchen Regeln in der Nutzung eingeschränkt?
- 2) Sollte das Netz an dieser Stelle ausgebaut werden, um Zeiten mit Engpässen zu reduzieren? Wann lohnt sich eine Investition in die Erweiterung des Netzes?

Beide Fragenkomplexe – die zum Engpassmanagement und die zur Netzausbauentscheidung – lassen sich analysieren, wenn man das Konzept des externen Effektes hinzuzieht: Eine Person, die ein Netz nutzt, bei dem Engpässe auftreten, behindert andere Personen, die dieses Netz nutzen oder nutzen wollen. Wenn diese Behinderung nicht in das Kalkül der Person eingeht (wer denkt, wenn er in einen Stau fährt, an die dahinter fahrenden Personen?), dann ist dies ein externer Effekt – d.h. eine Wirkung außerhalb des eigenen Kalküls.

Es ist eine alte Erkenntnis aus der Ökonomie – entwickelt von Pigou (1920) –, dass man zu einer effizienten Allokation, bzw. in unserem Fall zu einer effizienten Netznutzung, kommen kann, indem man diesen externen Effekt durch eigene Preise in der Entscheidungsfindung berücksichtigt, sprich internalisiert. CO₂-Preise sind das bekannteste Beispiel dafür. Der negative Effekt auf das Klima mit den damit einhergehenden menschlichen und wirtschaftlichen Schäden durch die Emission von CO₂ wird durch solche Preise internalisiert. So müssen beispielsweise Kohlekraftwerke, welche mehr CO₂ emittieren als Gaskraftwerke, mehr zahlen und werden dadurch weniger wettbewerbsfähig. In Großbritannien wurde mit Hilfe solcher CO₂-Preise die Stromproduktion durch Kohlekraftwerke fast vollständig zurückgefahren (Reuters, 2018).

Eine solche Logik greift auch bei Engpässen in Netzen und den dabei auftretenden externen Effekten. Eine Person, die in einen Stau hineinfährt, berücksichtigt bei einer verkehrsaufkommensabhängigen Abgabe dann nicht nur die individuellen Fahrtkosten und die Kosten der eigenen Zeit, die sie im Stau verbringt, sondern auch die Erhöhung der Kosten, die den hinter ihr fahrenden Personen entstehen. Manche Personen werden dann

auf die Fahrt verzichten oder Alternativen wie eine andere Route oder die Nutzung weiterer Transportmittel wie den öffentlichen Nahverkehr wählen. Insofern ist der Begriff „Anti-Stau-Maut“ zutreffend. Wenn diese richtig gewählt wird, verhindert sie den Stau, wie Beispiele aus den USA und Singapur zeigen (Mietsch, 2007).

Die Bepreisung des externen Effektes dient also dazu, das Engpassmanagement effizienter zu gestalten. Sie hilft aber auch dabei, die Netzausbauentscheidungen zu steuern. Wenn solche Preise vorliegen, wird das Netz in Peak-Zeiten unmittelbar weniger beansprucht, da Nachfrager auf die Nutzung des Netzes verzichten, um die Bepreisung zu vermeiden. In einer dynamischen Perspektive wird dieser Effekt noch verstärkt, da sich durch die Bepreisung auch Anreize für Investitionsentscheidungen in den jeweiligen Märkten ergeben. Wenn etwa bei regionalen Strompreisen der Strom im Süden teurer wird, da dieser vermehrt im Norden produziert wird und das Stromnetz es nicht schafft, ihn nach Süden zu transportieren, werden Anreize geschaffen, Strom im Süden zu produzieren bzw. Produktion (also Verbrauchsstätten) im Norden anzusiedeln. Der statische und der dynamische Effekt reduzieren die Notwendigkeit eines Netzausbaus. Hinzu kommt, dass die Zahlungen für diese externen Effekte auch ein Signal darstellen, wo die Engpässe besonders schmerzhaft sind. So sollte ein Autobahnabschnitt umso eher ausgebaut werden, desto höher dort das Aufkommen der Anti-Stau-Maut ist.

Der optimale Ausbau bestimmt sich in einem aufwändigeren Kalkül, das die Höhe der Bepreisung des externen Effektes und die Kosten des Ausbaus berücksichtigt (Mohring & Harwitz, 1962). Grundsätzlich sollte so weit ausgebaut werden, bis die (marginale) Reduktion der externen Kosten gleich hoch ist wie die zusätzlichen (marginalen) Ausbaukosten.

3. Maßnahmen zur stärkeren Nutzersteuerung bei Netzengpässen

3.1 Verkehr

Der Straßenverkehr ist das Lehrbuchbeispiel für die effizienzsteigernde Wirkung einer auslastungsabhängigen Straßenmaut. Eine Anti-Stau-Maut, d.h. eine Nutzungsgebühr abhängig vom Verkehrsaufkommen, würde helfen, dass die Straße effizienter genutzt wird. Bei Parkplätzen ist dies offensichtlich: In allen größeren Städten ist es heute üblich, dass man zahlen muss, wenn man den öffentlichen Raum als „Parkplatz“ nutzt, da dieser knapp ist. Und am Wochenende und nachts wird es billiger, weil dann weniger Bedarf für Parkplätze ist.

In vielen Ländern wie beispielsweise Großbritannien, den USA oder Singapur wurde eine solche verkehrsabhängige Straßenmaut auch schon mit Erfolg umgesetzt (s. Mietsch, 2007). So konnte dort das Verkehrsaufkommen durch unterschiedliche Mautgebühren abhängig vom Wochentag und der Tageszeit gesenkt werden und die Durchschnittsgeschwindigkeit auf den Mautstraßen stieg an. Außerdem kann eine Straßenmaut insbesondere für Städte die Luft- und Lebensqualität verbessern und einen Standortvorteil bieten. Die Einnahmen aus der Maut können darüber hinaus helfen, die Verkehrsinfrastruktur besser zu erhalten. In Deutschland gibt es eine solche Anti-Stau-Maut bislang nicht. Es verwundert, dass wir uns längst daran gewöhnt haben, Gebühren für (knappen) Parkraum zu zahlen, aber Gebühren für (knappen) Straßenraum häufig kategorisch ausschließen.

In Deutschland verursachten Staus 2017 einen volkswirtschaftlichen Schaden von rund 80 Mrd. Euro (INRIX, 2018). Dazu zählen neben den Zeitkosten der Menschen, die in den Staus unnütz warten müssen, auch die Lärm- und Feinstaubbelastung vor allem in den Städten, die bei Stau noch größer ist als bei fließendem Verkehr. So verbrauchen Autos im Stau bis zu 80 % mehr Kraftstoff als im freien Verkehr. Zudem verursachen Staus außerhalb des Verkehrs Kosten: Sie führen zu erheblichen, nicht vorhergesehenen externen Effekten wie verringerte Risikoaversion, erhöhte Ungeduld und eine geringere Bereitschaft, anderen zu helfen.² Moderne GPS Systeme mit Stauumfahrung bieten zwar insofern Erleichterung, dass Fahrer vor Nutzung einer Straße über mögliche Staus informiert werden und deshalb ausweichen können, sie machen aber auch das Fahren mit dem Auto attraktiver, was zu einem erhöhten Verkehrsaufkommen nicht nur auf den Hauptverkehrsadern, sondern auch auf Land- und Seitenstraßen von Wohngebieten beiträgt (Cramton et al., 2018).

Differenzierter als eine auslastungsabhängige Straßenmaut ist die sogenannte smarte Maut, wie es sie etwa in Kalifornien gibt (Mietsch, 2007). Der US-Bundesstaat erhebt auf einer vierspurigen Straße nur für die linke Spur eine Maut. Wenn sich ein Stau bildet, wird für die Nutzung der linken Spur eine Gebühr verlangt. Sie ist flexibel und wird je nach Verkehrsaufkommen so eingestellt, dass der Verkehr auf dieser Spur wieder zügig vorankommt: Ist sie zu niedrig, wollen zu viele auf die linke Spur, und dann gibt es auch dort Stau. Dann wird die Gebühr angehoben. Ist sie zu hoch, fährt kaum ein Fahrzeug auf der linken Spur. Dann wird die Gebühr gesenkt, damit mehr Fahrzeuge dort fahren, aber nur so stark, dass

2 Vorläufige Ergebnisse aus Werthschulte et al. (2021).

der Verkehrsfluss auf der linken Spur immer noch zügig ist. Wenn die Maut richtig eingestellt ist, gewinnen am Ende alle: Die, die ganz links fahren und die Gebühr zahlen, kommen schneller voran (und haben sich ja bewusst dafür entschieden, diese Gebühr zu zahlen), und diejenigen, die auf den drei rechten Spuren verbleiben, stellen sich auch besser, weil dort das Fahrzeugaufkommen und damit auch der Stau geringer wird. Denn der zügige Verkehrsfluss auf der linken Spur sorgt dafür, dass insgesamt mehr Autos durch den Stau kommen, und damit auch auf den rechten Spuren weniger Autos unterwegs sind. Darüber hinaus zahlen Autos mit mindestens drei Personen weniger als leere Autos, um Anreize zu schaffen gemeinsam zu fahren und das Verkehrsaufkommen umso mehr zu verringern. Eine solche smarte Maut, die das Verkehrsgeschehen und die lokalen Emissionen berücksichtigen würde, wäre ein echter Schritt vorwärts.

Cramton et al. (2018) gehen sogar noch einen Schritt weiter und schlagen vor, über sogenanntes Dynamic Pricing (Dynamisches Bepreisen) und GPS-Ortung aller Autos die Straßennutzung zu überwachen und gegebenenfalls in Echtzeit über Straßennutzungsgebühren einzugreifen, um Staus zu vermeiden, bevor sie sich bilden. Ein solches System steckt in vielerlei Hinsicht noch in den Kinderschuhen, zeigt jedoch die Möglichkeiten auf, die es zukünftig geben könnte. Sorge bereitet insbesondere die Gewährleistung von Gleichheit und Privatsphäre im System, da möglicherweise ärmere Bevölkerungsschichten stärker von der Straße verdrängt werden. Hier bedarf es geeigneter Ausgleichsmodalitäten. Außerdem sind die technischen Möglichkeiten noch nicht ausgereift, um alle Autos zu geringen Kosten genau zu orten und anschließend in entsprechenden Modellen optimale Preise in Echtzeit zu ermitteln.

Nicht nur in Deutschland ist die Skepsis der Politik sowie der Wählerschaft bezüglich des Nutzens einer Maut recht hoch. Allerdings zeigt das Beispiel Stockholms, dass sich diese Skepsis bei guter Umsetzung in Zustimmung wandeln kann (s. Mietsch, 2007). In Stockholm haben sich vor Einführung der City-Maut 69 % der Bevölkerung gegen eine Gebührenerhebung ausgesprochen, während achteinhalb Monate später 51 % der Wähler ihre Zustimmung dafür ausdrückten.

In Deutschland war auf politischer Ebene die Einführung einer „PKW-Maut für Ausländer“ lange Zeit Projekt der CSU. Beabsichtigt war, eine Straßenbenutzungsgebühr für Personenkraftwagen einzuführen, und gleichzeitig die Kraftfahrzeugsteuer zu reduzieren. Damit wäre die finanzielle Belastung primär bei den ausländischen Nutzern deutscher Autobahnen angefallen. Der Europäische Gerichtshof stoppte diese Koppelung.

Nach dieser Erfahrung wurde das grundsätzliche Mittel einer Maut zur effizienten Nutzung der Straße aus dem Fokus genommen. Zwar befür-

worten Bündnis 90/Die Grünen sowie Die Linke eine strengere Regulierung des LKW-Verkehrs durch die LKW-Maut in ihren Wahlprogrammen. Eine PKW-Maut wurde aber nur in Programmen von Kleinparteien diskutiert. Hierbei fordert Volt als einzige Partei ausdrücklich die Überprüfung von Kosten und Nutzen einer PKW-Maut gegenüber der Erhöhung von Energie- und PKW-Steuern. Auf der anderen Seite äußern sich lediglich die Freien Wähler und die Piraten-Partei gegen die Einführung einer PKW-Maut, u.a. wegen der Sorge vor der Überwachung durch etwaige Mauterfassungssysteme.

Dabei findet die Option einer City-Maut in der Wissenschaft viel Zustimmung: Neben dem Wissenschaftlichen Beirat beim BMWi (s. May, 2019; Wissenschaftlicher Beirat BMWi, 2019), sprechen sich auch Wissenschaftler des ZEW (s. Achtnicht et al., 2018) sowie des RWI und der Mercator-Stiftung (s. RWI-Leibniz-Institut für Wirtschaftsforschung & Stiftung Mercator, 2019) dafür aus. So schafft eine City-Maut im Gegensatz zu Diesel-Fahrverboten eine Wahlmöglichkeit für Autofahrer und ist deshalb laut ZEW zu präferieren. Auch der Deutsche Städtetag hat sich 2019 für die Erprobung eines solchen Systems in Städten, die dies wünschen, ausgesprochen (Spiegel, 2019). Eine flächendeckende Umsetzung wurde jedoch trotz vielversprechender Beispiele aus dem Ausland abgelehnt. Auch der ADAC äußerte sich im Falle von Berlin und Brandenburg ablehnend. In Deutschland wurde bzw. wird eine City-Maut in manchen Städten wie etwa Hamburg, Berlin oder München geprüft und diskutiert, während sie in Düsseldorf oder Frankfurt/Main bereits abgelehnt wurde.

Die Anti-Stau-Maut mit ihrer Steuerungswirkung bei knappen öffentlichen Gütern lässt sich auch auf andere Verkehrsbereiche übertragen. Das damalige Gutachten empfahl den Handel mit Start- und Landerechten an Flughäfen, den die Monopolkommission (2016) in ihrem XXI. Hauptgutachten im Kapitel zur Flughafenregulierung aufgegriffen hat. Auch Bahntrassen sind knapp. Wie die Monopolkommission berichtet, betrug laut Angaben der Europäischen Kommission die Gesamtlänge überlasteter Schienenwege im Jahr 2018 in Deutschland insgesamt 721 Trassenkilometer; Platz zwei der meist überlasteten Schienenwege im europäischen Vergleich. Auch hier würde für eine effiziente Zuteilung dieser Trassen der Einsatz von Knappheitszuschlägen helfen, die zwar im Eisenbahnregulierungsgesetz vorgesehen sind, aber derzeit nicht angewendet werden (Monopolkommission, 2021, Fn. 183).

3.2 Elektrizität

Deutschland bildet, gemeinsam mit Luxemburg, eine Preiszone im Strommarkt. Der Großhandels-Strompreis ist in ganz Deutschland gleich. Das bedeutet, dass ein Unternehmen im Süden Deutschlands genauso gut Strom von einem Erzeuger im Süden Deutschlands wie von einem im Norden kaufen kann.

Allerdings ist das Stromnetz in Deutschland nicht ausreichend ausgebaut, um diesen Handel dann auch physisch umzusetzen. Bei Netzengpässen etwa zwischen dem Norden und dem Süden Deutschlands muss der Netzbetreiber einem Erzeuger im Norden untersagen, Strom zu produzieren, und einem weiteren Erzeuger in Süddeutschland anordnen, Strom in den süddeutschen Teil des Netzes einzuspeisen. Dafür erhalten die Erzeuger Kompensationen, die sogenannten Redispatch-Aufwendungen. Diese Redispatch-Kosten, die in den Jahren 2007 bis 2011 stets weniger als 50 Millionen Euro betrugen, stiegen im Jahr 2015 auf über 400 Millionen Euro. 2019 lagen sie bei 207 Millionen (Fekete, 2020, S. 10).

Die Netzbetreiber mussten also immer häufiger Redispatch-Maßnahmen vornehmen, und es ist davon auszugehen, dass mit der zunehmenden Stromeinspeisung durch erneuerbare Energien, die besonders im Norden Deutschlands angesiedelt werden, diese noch zunehmen werden.

Diese Redispatchkosten werden mit den Netzegebühren den Stromnutzern in Rechnung gestellt. Dabei zahlen deutsche Haushalte schon heute die höchsten Strompreise weltweit (Global Petrol Prices, 2021). So lag gemäß dieser Studie der Preis für eine Kilowattstunde im März 2021 in Deutschland bei 0,37 US-Dollar, in Frankreich hingegen bei 0,21.

Die Engpässe im Stromnetz, die die Redispatchaufwendungen verursachen, werden noch lange bestehen bleiben, da der Stromnetzausbau stockt. Seit 2009 wurden mehr als 7.500 Kilometer Netzausbauvorhaben beschlossen, von denen jedoch erst 950 Kilometer, also knapp 13 Prozent, in Betrieb sind. Der größte Engpass liegt beim Transport von Norden nach Süden: Wegen der starken Erzeugung im Norden und der großen Nachfrage im Süden würde es deshalb helfen, wenn mehr Strom im Norden verbraucht werden würde und weniger im Süden, und wenn sich Stromerzeuger eher im Süden als im Norden ansiedeln würden.

Die EU-Kommission, die für den Stromhandel im europäischen Binnenmarkt zuständig ist, hätte Deutschland gerne schon seit 2016 in zwei Strompreiszonen aufgeteilt, um dem Problem der Durchleitung von Nord nach Süd beizukommen. Zwei Preiszonen würden tendenziell dazu führen, dass der Strom im Norden billiger wäre als im Süden. Windstrombetreiber hätten mehr Anreiz, sich im Süden anzusiedeln, selbst wenn dort

weniger Wind weht, weil dort die Preise höher wären. Und Unternehmen hätten einen Anreiz, sich im Norden niederzulassen, wo die Strompreise günstiger wären, was auch sinnvoll wäre, weil dort mehr Strom erzeugt wird. Der Energiesektor wäre effizienter, und der Netzausbau wäre weniger aufwendig.

In der Diskussion anlässlich der Blackout-Gefahr im Juni 2020 wurde erneut Kritik am bestehenden System geübt (Janzing, 2020). Diesbezüglich würde die Einführung mehrerer Preiszonen auch die Resilienz gegenüber einem Blackout erhöhen. In den Wahlprogrammen für die Bundestagswahl 2021 kam das genannte Zwei-Preis-zonen-System allerdings bei keiner Partei zur Sprache.³

Zwei Preiszonen sind auch nur ein recht grobes Instrument, um die Engpässe im deutschen Stromnetz zu internalisieren. Zielgenauer sind die sogenannten Knotenpreise (Weyer & Müsgens, 2020): Dazu wird an jedem Standort (dem „Knoten“) ermittelt, inwiefern Stromnachfrage und Stromangebot dort dazu beitragen, die Netzengpassproblematik zu mindern. Dort, wo reichlich Strom vorhanden ist, sinken dann die (Knoten-)Preise, dort, wo wenig Strom vorhanden aber eine hohe Stromnachfrage herrscht, steigen die Preise. Solche Knotenpreise gibt es u.a. in den USA, wo sie aufgrund der vergleichsweise schwächeren Netzinfrastuktur bereits vor mehr als zwanzig Jahren eingeführt wurden. Mit den Knotenpreisen bekommt Knappheit – in diesem Fall knappe Stromleitungskapazität – einen Preis. Wie oben bereits dargestellt wird das Netz, trotz Ausbau, aufgrund der steigenden Belastung durch zunehmende Einspeisung volatiler erneuerbarer Energien mit stärker lokal differenzierter Erzeugung, immer öfter an seine Grenzen kommen. Damit steigt die Bedeutung örtlich differenzierter Preise zur lokalen Steuerung des Angebots und der in den Strommarkt zu integrierenden flexiblen Nachfrage.

Knotenpreise könnten in das derzeitige System eines Großhandelsmarktes über die schon am Anfang dieses Textes erwähnte verallgemeinerte Merit Order integriert werden. Da diese doch recht technisch ist, wurde sie im Gutachten in einer eigenen Erklärbox dargestellt, die hier wörtlich übernommen wird.

3 Im September 2022 wurde eine Aufteilung Deutschlands in unterschiedliche Strompreiszonen von Regierungsvertretern von Niedersachsen, Schleswig-Holstein und Mecklenburg-Vorpommern ins Gespräch gebracht.

Verallgemeinerte Merit Order:

Im Rahmen einer disaggregierten, knotenbasierten Bepreisung kann die verallgemeinerte Merit Order wie folgt abgeleitet werden (Knieps, 2013, S. 153ff.). Es wird davon ausgegangen, dass für das ganze Netzgebiet ein einziger wettbewerblicher Großhandelsmarkt (Strombörse) für Elektrizität existiert. Der Netzbetreiber erhebt knotenbasierte Einspeise- und Ausspeisetarife auf der Basis der Opportunitätskosten der Netzinanspruchnahme. Diese bestehen aus Systemexternalitäten durch Stromverluste und veränderte Netzauslastung. Die dezentralen Stromerzeuger speisen gegen Zahlung der Netzgebühr Strom in das Netz ein. Die verallgemeinerte Merit Order gibt an, an welchen Netzknoten sich eine Netzeinspeisung lohnt, so dass die Erzeugungskosten und der Einspeisetarif die marginale Zahlungsbereitschaft auf dem Großhandelsmarkt nicht übersteigt. Die Knotenpreise an den Ausspeiseknoten reflektieren die Summe des (einheitlichen) Großhandelspreises und die knotenabhängigen Ausspeisetarife. Der Stromhandel wird über eine Börse abgewickelt, wobei das gesamte Stromangebot und die gesamte Großhandelsnachfrage innerhalb des Netzgebietes (mittels einer einzigen Energieausgleichsbedingung) ermittelt werden. Knotenbasierte Einspeise- und Ausspeisetarife führen folglich nicht zu einer Marktaufteilung („Market Splitting“) mit unterschiedlichen Zonen; insbesondere hängt aufgrund der verallgemeinerten Merit Order der Großhandelspreis nicht davon ab, an welchen Knoten im Netz eingespeist wird. Entscheidend sind die Summe von Grenzkosten der Erzeugung und der Netzeinspeisepreis an den jeweiligen Knoten. Allerdings müssen die Nachfrager, abhängig von ihrem Standort, zusätzlich zu dem einheitlichen Großhandelspreis die Netzausspeisegebühren in Höhe der Opportunitätskosten der Netzausspeisung bezahlen. Arbitragemöglichkeiten entstehen hierdurch nicht, da die Nachfrager den Strom nur an ihrem Standort konsumieren können.

Quelle: BMWi (2014, S. 9).

Eine Umstellung auf Knotenpreise birgt auch Nachteile in sich. Gegen die Einführung eines Knotenpreissystems spricht der hohe Aufwand, sofern die Verteilernetzebene miteinbezogen werden soll, sowie die grundlegende Umgestaltung des vorherrschenden Marktdesigns (s. Weyer & Müsgens, 2020). So kann die neue räumliche Marktabgrenzung zu mehr Marktmacht führen und müsste entsprechend kontrolliert werden. Außerdem benötigt das Knotenpreissystem eine einheitliche Netzbetriebsführung, die insbesondere in Grenzgebieten und dort, wo mehrere Betreiber Netze besitzen, schwer umsetzbar ist.

Dennoch ist zu erwarten, dass mittelfristig regionale Preise – wie die verallgemeinerte Merit Order oder Knotenpreise – auch in Deutschland

eingesetzt werden. So sprechen sich die Grünen in ihrem Fraktionsbeschluss vom April 2021 (Bündnis 90/Die Grünen Bundestagsfraktion, 2021) dafür aus, „perspektivisch die Einführung von Preiszonen oder von Knotenpreisen vorzubereiten“. In einer Frage an die Bundesregierung im Juni dieses Jahres stellt die FDP einen Umstieg auf ein Knotenpreissystem als eine passende Lösung für Netzengpässe dar.

3.3 Telekommunikation und Internet

Die Diskussion um das Management von Engpässen beim Transport von Datenpaketen im Internet wurde damals im Jahr 2014 unter dem Begriff der Netzneutralität geführt. Das Gutachten hielt fest: „Der Umgang mit Datenpaketen im Internet sowie die mögliche Priorisierung von einzelnen Datenpaketen werden unter dem Leitbegriff der „Netzneutralität“ diskutiert. Bei der Netzneutralitätsdebatte geht es um die Frage, ob und in welchem Maße Anbieter von Internet-Zugangsdiensten ein aktives Verkehrsmanagement der Datenpaketübertragung betreiben dürfen oder ob dieses Verkehrsmanagement durch Regulierung eingeschränkt werden soll. In einer extremen Ausprägung könnten Netzbetreiber regulatorisch verpflichtet werden, allen Datenpaketen die gleiche Priorität innerhalb ihres Netzes zuzuordnen. Auch geht es um die Frage, ob Netzbetreiber im Rahmen ihres Verkehrsmanagements bei Anbietern von Internet-Anwendungsdiensten wie beispielsweise Google, Amazon oder Telekom Entertainment Preis- und Qualitätsdifferenzierungen vornehmen dürfen.“

Der Wissenschaftliche Beirat hatte sich damals auf die Seite derer geschlagen, die sich für Preis- und Qualitätsdifferenzierungen ausgesprochen hatten. So könnten Engpässe im Internet bepreist werden und insbesondere Produkte mit unterschiedlichen Qualitäten bei der Durchleitung von Datenpaketen angeboten werden. Die Empfehlung lautete: „Im Telekommunikationsnetz sollen Qualitätsunterschiede in der Datenübertragung nicht beschränkt werden. Eine damit einhergehende Preisdifferenzierung von Anwendungsdiensten hinsichtlich ihrer Übertragungsqualität ermöglicht neue Geschäftsmodelle und Dienstleistungen.“

2014 war die Diskussion zur Netzneutralitätsregulierung im vollen Gang. Die „Netzneutralitätsverordnung“ der EU (VO 2015/2120/EU) trat im November 2015 in Kraft (Europäische Union, 2015). Auch wenn die Schlagzeilen damals betonten, dass dadurch „Netzneutralität“ gesichert werde, hat die Verordnung doch auch Möglichkeiten geschaffen, qualitätsgesicherte Dienste anzubieten (BMW, 2021). Ein gewisses Management des Datenverkehrs ist erlaubt, es muss aber transparent, diskriminierungs-

frei und verhältnismäßig durchgeführt werden und darf keine kommerziellen Interessen verfolgen (Art. 3 Abs. 3 VO 2015/2120/EU). Die Netzneutralität darf im Sonderfall einer zeitweisen Gefahr einer Netzüberlastung oder zur Erfüllung rechtlicher Auflagen eingeschränkt werden. Solange die Kapazität der Leitungen ausreichend ist, dürfen besondere Dienste angeboten werden für die Abweichungen von den Regeln gelten (Art. 3 Abs. 5 VO 2015/2120/EU). Diese Dienste dürfen aber Endnutzer nicht einschränken.

Aus ordnungspolitischer Sicht steht eine solche Regulierung von Datenverkehrsdiensten allerdings nicht nur zum Prinzip liberalisierter Dienstemärkte im Widerspruch, sondern stellt auch eine Überregulierung dar, die die unternehmerische Gestaltung von Suchprozessen unterbindet (s. Knieps & Stocker, 2015).

Die Verordnung bekam im September 2020 besonders große Brisanz, als eine Klage gegen einen ungarischen Netzanbieter direkten Einfluss auf viele Mobilfunknutzer hatte (s. Hauck & Martin-Jung, 2017; Kornmeier, 2020; Wilkens, 2020). Dabei ging es um die sogenannte Zero-Rating-Option, die Teil vieler Mobilfunkverträge ist. Diese erlaubt, dass die Nutzung bestimmter Apps, wie beispielsweise WhatsApp, Instagram oder Spotify, nicht auf das vertragliche Datenvolumen angerechnet wird und somit unbegrenzt genutzt werden kann, selbst wenn das vertragliche Volumen bereits erschöpft ist.

In einem ersten Urteil stellte der EuGH zunächst am 15. September 2020 fest, dass es unzulässig ist, nach Ausschöpfung des gebuchten Datenvolumens die über den Zero-Rating-Tarif gebuchten Anwendungen weiterhin ungedrosselt zu ermöglichen und übrige Anwendungen zu blockieren oder zu verlangsamen.⁴ Ob über diesen speziellen Fall hinaus Zero-Rating-Angebote zulässig sind, war in der Folge umstritten. Ziel des Urteils ist die Wahrung der Netzneutralität durch das Verbot von Diskriminierung bzw. Bevorteilung einzelner Dienste.

Trotz dieses Urteils werden in Deutschland weiterhin diverse Zero-Rating-Tarife von den Mobilfunkunternehmen wie Telekom's „Stream On“ oder Vodafone's „Vodafone Pass“ angeboten. Damit die Datenmengen in ihrem Netzwerk nicht zu groß wurden drosselte beispielsweise die Telekom 2017 in ihrem Netzwerk den Download beim Streaming von Filmen, indem ausschließlich eine niedrigere Bildqualität zugelassen wurde (Zeit, 2021). Vodafone passte ebenfalls die Allgemeinen Geschäftsbe-

4 EuGH, Urteil vom 15.9.2020 – Az. C-807/18, C-39/19, ECLI:EU:C:2020:708, Rn. 54.

dingungen des Vertrags entsprechend an, um diese Möglichkeit nutzen zu können. Außerdem ließen die Verträge die uneingeschränkte Datennutzung im europäischen Ausland nicht zu, obwohl hier nach „Roam like at Home“ keine Unterscheidung vorliegen darf. Zudem waren Klauseln in den Verträgen von Vodafone enthalten, die eine Nichtanrechnung von Datenvolumen ausschlossen, wenn eine Nutzung der Dienste mittels Tethering (Nutzung des Smartphones als Hotspot) erfolgte. Gegen diese Klauseln ging einerseits die Bundesnetzagentur mittels einer Anordnung vor und andererseits die Verbraucherzentrale Bundesverband (VZBV) auf dem Zivilrechtsweg. Die Unternehmen traten diesen Änderungsverlangen im gerichtlichen Verfahren entgegen. Die Gerichtsverfahren dauern noch immer an. Jedoch urteilte der EuGH im September 2021 auf Anfrage der sich damit befassenden deutschen Gerichte, dass die Verträge der beiden Unternehmen rechtswidrig sind (EuGH, 2021).⁵ Deshalb wird derzeit nicht mehr davon ausgegangen, dass die Unternehmen ihre Klagen vor den deutschen Gerichten gewinnen können. Die VZBV erfreute das Urteil und sie sieht sich in ihrem Verständnis der Verordnung bestätigt, dass Diskriminierung im Datenverkehr nicht auf Grundlage von kommerziellen Erwägungen geschehen darf (Rudl, 2021).

Die Verfahren zeigen, dass auf der einen Seite der Bedarf für besonders datenintensive Optionen gegeben ist, diese aber auf der anderen Seite stark reguliert werden. Die Urteile tragen dazu bei, dass Mobilfunkanbieter zum Beispiel Tier-Videos in HD gleichwertig übermitteln müssen wie Daten-Verkehr zwischen Unternehmen. Somit wird ihnen die Möglichkeit genommen, ihre Datenleitungen effizient auf ihre Kunden zu verteilen und die Gefahr von Überlastung und Drosselungen erhöht sich. Eine Änderung der Rechtslage scheint aufgrund der kürzlich ergangenen Urteile des EuGH zu den deutschen Zero-Rating-Gerichtsverfahren, in denen er sein Urteil im Fall des ungarischen Anbieters bestätigte und dessen Reichweite klarstellte, äußerst unwahrscheinlich.

Günter Knieps selbst sieht über die Frage der Netzneutralität hinaus noch großen Handlungsbedarf bei der Regulierung der Telekommunikationsnetze. In seiner Abschiedsvorlesung an der Wirtschafts- und Verhaltenswissenschaftlichen Fakultät der Albert-Ludwigs-Universität Freiburg wies er darauf hin, dass das Internet der Dinge zu fundamentalen Herausforderungen für die traditionellen Netzindustrien führt, da physische

5 S. EuGH, Urteil vom 2.9.2021 – Az. C-854/19, E-CLI:EU:C:2021:675; EuGH, Urteil vom 2.9.2021 – Az. C-34/20, E-CLI:EU:C:2021:677; EuGH, Urteil vom 2.9.2021 – Az. C-5/20, E-CLI:EU:C:2021:676.

Netzdienstleistungen immer stärker auf Echtzeitübermittlung basieren und dadurch adaptive und standortsensitive Informationen bereitgestellt werden müssen. Die vielfältigen Potenziale des Internets der Dinge zu nutzen und gleichzeitig die Gefahren dieser Technologie in den Griff zu bekommen, sei eine gesellschaftspolitische Herausforderung. Dafür brauche es einen interdisziplinären Ansatz, der weit über das Gebiet der Netzökonomie hinausgeht.

4. Fazit

Das Fazit im damaligen Gutachten des Beirats beim BMWi war sehr kurz gehalten und passte auf eine halbe Seite: Es empfahl die Prüfung der „Einführung einer auslastungsabhängigen Straßenmaut für LKW und PKW“, die Schaffung der „Voraussetzungen für die Verwendung auslastungsabhängiger Netznutzungsentgelte“ im Stromnetz, und dass „die Qualitätsunterschiede in der Datenübertragung [im Telekommunikationsnetz] nicht beschränkt werden.“ Das Gutachten war seiner Zeit voraus. Heute werden die Instrumente einer intelligenten Straßenmaut und regionaler Strompreise auf vielen Ebenen diskutiert – in der Wissenschaft und in der Politik. Und auch die Probleme einer zu eng ausgelegten Netzneutralität im Internet werden immer deutlicher. Gute Ökonomie, zu deren Protagonisten Günter Knieps zählt, wird sich am Ende durchsetzen.

Referenzen

- Achtnicht, M., Kesternich, M., & Sturm, B. (2018). *Die „Diesel-Debatte“: Ökonomische Handlungsempfehlungen an die Politik*. Wirtschaftsdienst, 98(8), 574–577. <https://doi.org/10.1007/s10273-018-2333-4>.
- Bild (2014, 18. November). Extra-Gebühr in Stoßzeiten – Regierungs-Experten fordern Stau-Maut auf Autobahnen. *Bild*. <https://www.bild.de/geld/wirtschaft/experten/fordern-stau-maut-auf-autobahnen-38608474.bild.html>.
- BMWi (2014). *Engpassbasierte Nutzerfinanzierung und Infrastrukturinvestitionen in Netzsektoren—Gutachten des Wissenschaftlichen Beirats beim Bundesministerium für Wirtschaft und Energie*. Bundesministerium für Wirtschaft und Energie. https://www.bmwi.de/Redaktion/DE/Publikationen/Ministerium/Veroeffentlichung-Wissenschaftlicher-Beirat/wissenschaftlicher-beirat-engpassbasierte-nutzerfinanzierung-und-infrastrukturinvestitionen-in-netzsektoren.pdf?__blob=publicationFile&v=4.
- BMWi (2021). Netzneutralität. *Bundesministerium für Wirtschaft und Energie*. <https://www.bmwi.de/Redaktion/DE/Artikel/Digitale-Welt/netzneutralitaet.html>.

- Bündnis 90/Die Grünen Bundestagsfraktion (2021). *Strommarktdesign für eine sichere und kostengünstige Stromversorgung auf dem Weg hin zu 100 % Erneuerbaren Energien* (Fraktionsbeschluss vom 20. April 2021). <https://www.gruene-bundestag.de/files/beschlusse/beschluss-strommarktdesign.pdf>.
- Cramton, P., Geddes, R. R., & Ockenfels, A. (2018). Set road charges in real time to ease traffic. *Nature*, 560, 23–25. <https://doi.org/10.1038/d41586-018-05836-0>.
- EuGH (2021). „Nulltarif-Optionen“ verstoßen gegen die Verordnung über den Zugang zum offenen Internet (Pressemitteilung Nr. 145/21). Gerichtshof der Europäischen Union. <https://curia.europa.eu/jcms/upload/docs/application/pdf/2021-09/cp210145de.pdf>.
- Europäische Union (2015). Verordnung (EU) 2015/2120 des Europäischen Parlaments und des Rates vom 25. November 2015 über Maßnahmen zum Zugang zum offenen Internet und zur Änderung der Richtlinie 2002/22/EG über den Universaldienst und Nutzerrechte bei elektronischen Kommunikationsnetzen und -diensten sowie der Verordnung (EU) Nr. 531/2012 über das Roaming in öffentlichen Mobilfunknetzen in der Union. <https://eur-lex.europa.eu/legal-content/DE/ALL/?uri=celex%3A32015R2120>.
- Fekete, P. (2020). *Redispatch in Deutschland. Auswertung der Transparenzdaten. April 2013 bis einschließlich September 2020*. Bundesverband der Energie- und Wasserwirtschaft e.V (BDEW). https://www.bdew.de/media/documents/2020_Q3_Bericht_Redispatch_GOQPSvY.pdf.
- Global Petrol Prices (2021). Strompreise. https://de.globalpetrolprices.com/electricity_prices/ [Abgerufen am 26.10.2021].
- Hauck, M., & Martin-Jung, H. (2017, 15. Dezember). Auch in Deutschland ist die Netzneutralität durchlöchert. *Süddeutsche Zeitung*. <https://www.sueddeutsche.de/digital/netzneutralitaet-in-deutschland-auch-in-deutschland-ist-die-netzneutralitaet-durchloechert-1.3791748>.
- INRIX (2018). Global traffic scorecard. <http://inrix.com/scorecard/> [Abgerufen am 26.10.2021].
- Janzing, B. (2020, 11. April). Wenn der Markt den Strom verknappt. *Taz am Wochenende*. <https://taz.de/Wenn-der-Markt-den-Strom-verknappt!/5676425/>.
- Knieps, G. (2013). Renewable energy, efficient electricity networks and sector-specific market power regulation. In F.P. Sioshansi (Hrsg.). *Evolution of Global Electricity Markets: New paradigms, new challenges, new approaches* (S. 147–168), Amsterdam et al.: Elsevier.
- Knieps, G., & Stocker, V. (2015). Network Neutrality Regulation: The Fallacies of Regulatory Market Splits. *Intereconomics* 50, 46–51.
- Kornmeier, C. (2020, 15. September). Keine App darf bevorzugt werden. *Tageschau*. <https://www.tagesschau.de/wirtschaft/eu-netzneutralitaet-faq-101.html>.
- May, F.R. (2019, 31. Mai). Mautpreise: Kein Patentrezept für Verkehrsprobleme in Innenstädten. *Handelsblatt*. <https://www.handelsblatt.com/politik/konjunktur/research-institute/hri-analyse-zu-strassengebuehren-mautpreise-kein-patentrezept-fuer-verkehrsprobleme-in-innenstaedten/24391114.html>.

- Mietsch, F. (2007). *City-Maut: Internationale Erfahrungen, Perspektiven für Deutschland*. Friedrich-Ebert-Stiftung. <http://library.fes.de/pdf-files/stabsabteilung/05057.pdf>.
- Monopolkommission (2016). *Wettbewerb 2016*. (21. Hauptgutachten der Monopolkommission gemäß § 44 Abs. 1 Satz 1 GWB.) https://www.monopolkommission.de/images/HG21/HGXXI_Gesamt.pdf.
- Monopolkommission (2021). *Bahn 2021: Wettbewerb in den Takt!* (8. Sektorgutachten. Gutachten der Monopolkommission gemäß § 78 ERegG.) https://www.monopolkommission.de/images/PDF/SG/8sg_bahn_volltext.pdf.
- Mohring, H. & Harwitz, M. (1962). *Highway Benefits: An Analytical Framework*. Evanston, IL, USA: Northwestern University Press.
- Pigou, A. (1920). *The Economics of Welfare*. London, UK: Palgrave Macmillan.
- Reuters (2018, 5. Januar). Britain outlines plans for 2025 coal-power phase out. *Reuters*. <https://www.reuters.com/article/uk-britain-coal-phase-out-idUKKBN1E U15P>.
- Rudl, T. (2021, September 2). Zero-Rating-Verbot: EuGH spricht Machtwort zu Netzneutralität. *Netzpolitik.org*. <https://netzpolitik.org/2021/zero-rating-verbot-eugh-spricht-machtwort-zu-netzneutralitaet/><https://netzpolitik.org/2021/zero-rating-verbot-eugh-spricht-machtwort-zu-netzneutralitaet/>.
- RWI-Leibniz-Institut für Wirtschaftsforschung & Stiftung Mercator (Hrsg.). (2019). *Weniger Staus, Staub und Gestank per sozial ausgewogener Städte-Maut: Gemeinsames Plädoyer initiiert von RWI – Leibniz-Institut für Wirtschaftsforschung und der Stiftung Mercator* (RWI Position No. 74). RWI – Leibniz-Institut für Wirtschaftsforschung. http://www.rwi-essen.de/media/content/pages/publikationen/rwi-positionen/rwi-position_74_plaedoyer_fuer_eine_staedtemaut.pdf.
- Spiegel (2019, 25. April). Kommunen sollen City-Maut testen dürfen. *Spiegel*. <https://www.spiegel.de/auto/aktuell/staedtetag-kommunen-sollen-city-maut-testen-duerfen-a-1264305.html>.
- Tiroler Tageszeitung (2014). Ökonomen fordern für Deutschland „Stau-Maut“ – Zeitung. *Tiroler Tageszeitung*, 17. November 2014. <https://www.tt.com/artikel/9276029/oekonomen-fordern-fuer-deutschland-stau-maut-zeitung>.
- Werthschulte, M., Löschel, A., Razzolini, L., & Price, M. (2021). The Hidden Costs of Traffic Congestion [Unveröffentlichtes Manuskript].
- Weyer, H., & Müsgens, F. (2020). *Netzeingpässe als Herausforderung für das Stromversorgungssystem: Regelungsfelder, Status quo und Handlungsoptionen* (Schriftenreihe Energiesysteme der Zukunft). https://energiesysteme-zukunft.de/fileadmin/user_upload/Publikationen/PDFs/ESYS_Analyse_Netzengpassmanagement.pdf.
- Wilkens, A. (2020, 15. September). EuGH: Zero Rating verletzt Netzneutralität. *Heise*. <https://www.heise.de/news/EuGH-Zero-Rating-verletzt-Netzneutralitaet-4894216.html>.
- Wissenschaftlicher Beirat beim BMWi (2019). *Eine marktorientierte Verkehrswende*. https://www.bmwi.de/Redaktion/DE/Downloads/Wissenschaftlicher-Beirat/brief-marktorientierte-verkehrswende.pdf?__blob=publicationFile&v=6.

Zeit (2021, 2. September). Schlappe für Telekom und Vodafone vor dem EuGH. *Zeit*. <https://www.zeit.de/news/2021-09/02/eugh-einschraenkungen-bestimmter-handytarife-rechtswidrig>.

Der disaggregierte Regulierungsansatz und die preußischen Eisenbahnen im 19. Jahrhundert

Bernhard Wieland¹

Abstract

Die preußische Eisenbahngeschichte wird bisweilen als Beleg dafür angesehen, dass Wettbewerb im Eisenbahnwesen nicht funktionieren könne. Fremdling und Knieps (1993) haben diese These einer eingehenden Analyse unterzogen. Der vorliegende Beitrag ergänzt diese Analyse in zweierlei Hinsicht. Zum einen wird eine Schrift von David Hansemann aus dem Jahr 1837 erörtert, die überraschende Ähnlichkeiten zu Knieps' Ansatz der Disaggregierten Regulierung aufweist. Zum anderen wird auf die von Fremdling und Knieps vernachlässigte Rolle Bismarcks in der preußischen Eisenbahnpolitik eingegangen.

Keywords:

Railroads, 19th Century, Prussia, Regulation, Competition

1. Einleitung

Günter Knieps ist in erster Linie für seine Beiträge zur Regulierungstheorie bekannt. Sein Name ist vor allem mit dem Ansatz der Disaggregierten Regulierung verknüpft, den er seit den 80er Jahren des vorigen Jahrhunderts in einer Vielzahl von Veröffentlichungen entwickelt und ausformuliert hat.² Weniger bekannt ist, dass er sich auch intensiv mit wirtschaftshistorischen Studien zu einzelnen Branchen beschäftigt hat, vor allem der Telekommunikation und dem Eisenbahnwesen, das in diesem Beitrag im Vordergrund stehen soll.

Im Jahr 1993 veröffentlichten Fremdling und Knieps einen Aufsatz, in dem gefragt wurde, warum es in den 70er Jahren des 19. Jahrhunderts zu

1 Bernhard.Wieland@mailbox.tu-dresden.de. Der Verfasser dankt Prof. Brunekreeft für hilfreiche Kommentare zur ersten Fassung dieses Beitrags.

2 Lehrbuchmäßige Darstellungen finden sich in Knieps (2001) und Knieps (2007).

einer Verstaatlichung der preußischen Eisenbahnen kam, obwohl das damals noch gültige Preußische Eisenbahngesetz von 1837 und die regulatorische Diskussion der 1870er Jahre im Prinzip auch Lösungen ermöglicht hätten, die modernen Regulierungsansätzen schon sehr nahe kommen, so auch der disaggregierten Regulierung. Und in der Tat: Wer sich mit der Regulierung der preußischen Eisenbahnen im 19. Jahrhundert beschäftigt, staunt über die Modernität der Diskussion. Nicht einmal zwei Jahre nach Eröffnung der ersten deutschen Eisenbahnstrecke Nürnberg-Fürth im Jahr 1835, wurden bereits Vorschläge erörtert, die schon fast alle wesentlichen Elemente der gegenwärtigen Regulierungsdiskussion enthalten, wie etwa die Trennung von Netz und Betrieb, Ausschreibungen für monopolistische Bottlenecks, oder Tarifmodelle wie Wertstaffel, Entfernungstaffel oder Grenzkostenpreise. In einer erstaunlichen Schrift des Industriellen, Bankiers und Politikers David Hanseemann von 1837 wird Wettbewerb auf dem Netz und sogar die Möglichkeit einer Demsetz-Auktion zur Vergabe von Betriebsrechten erörtert (s.a. Wieland, 2018). Ähnlichkeiten mit Knieps' Ansatz der disaggregierten Regulierung sind erheblich. Vergleichbare Überlegungen gab es aber auch an anderer Stelle. Das bereits erwähnte Preußische Eisenbahngesetz von 1838 beispielsweise sah in § 27 Wettbewerb auf dem Netz explizit als Möglichkeit vor, auch wenn diese Möglichkeit in der Folge praktisch nie umgesetzt wurde (mit zwei marginalen Ausnahmen zur Durchleitung von Verkehr).

Trotzdem nahm die Entwicklung zunächst einen Verlauf, der mit disaggregierter Regulierung nur wenig zu tun hatte. Ab ca. 1840 kam es zu einem stürmischen Aufbau von privaten Eisenbahnnetzen durch voll vertikal integrierte private Unternehmen, denen sich später, ab 1848, auch staatliche oder vom Staat aufgekaufte Unternehmen hinzugesellten (Fremdling & Knieps, 1993). Von einer Trennung zwischen Netz und Betrieb oder Zugang von Dritten zum Netz konnte keine Rede sein. Als die Situation in den 1870er Jahren zu einer krisenhaften Zuspitzung kam, mit Überkapazitäten, zunehmender Kartellierung und einem ineffizienten Tarifchaos, wurden die alten Argumente aus den 30er Jahren wieder hervorgeholt. Dennoch wurde auch diesmal die disaggregierte Regulierung verworfen und stattdessen auf die vollständige Verstaatlichung der preußischen Eisenbahnen gesetzt.

Weshalb war dies so? Sagt die preußische Eisenbahngeschichte etwas über die Funktionsfähigkeit des Wettbewerbs im Eisenbahnwesen aus?

In einem Aufsatz von 1993 haben Fremdling und Knieps die zweite Frage verneint und die erste unter Rückgriff auf zwei Argumente beantwortet, die teils der normativen Theorie der Regulierung, teils der politischen Ökonomie entstammen. Aus der normativen Theorie stammt das Argu-

ment der asymmetrischen Informationsverteilung zwischen Regulierungsbehörde und reguliertem Unternehmen. Nach Fremdling und Knieps war die dominierende Meinung in den 1870er Jahren, dass aufgrund der Informationsasymmetrien eine effiziente Regulierung der Netzzugangsgebühren nicht möglich sei. Ein fairer Wettbewerb auf dem Netz sei deshalb illusorisch. Das zweite, politökonomische Argument besagt, dass der preußische Staat in der Nationalisierung der Bahnen eine lukrative Möglichkeit zur Aufbesserung der Staatsfinanzen sah. Angesichts der finanziellen Schwierigkeiten, in denen viele Bahnen in den 1870ern steckten, ist dieses Argument allerdings etwas zweifelhaft. Rentabel wurden die Bahnen für den Staat erst später.

Ziel des vorliegenden Beitrages ist eine zweifache Ergänzung der Analyse von Fremdling und Knieps. Zum einen soll das Augenmerk noch einmal auf die Diskussion in der ersten Hälfte des 19. Jahrhunderts gerichtet werden, insbesondere auf den Beitrag David Hansemanns, in Ergänzung der von Fremdling und Knieps überwiegend betrachteten Diskussion der 1870er Jahre. Zum anderen soll die Frage der Nationalisierung in den 1870–90ern stärker unter dem Blickwinkel der allgemeinen politischen Situation und dabei vor allem der besonderen Rolle Bismarcks (auch seiner persönlichen finanziellen Interessen) betrachtet werden (Stern, 2000). Fremdling und Knieps klammern diesen zweiten Aspekt gänzlich aus. Nach Ansicht des Verfassers fehlt aber damit der deutschen Eisenbahngeschichte dieser Zeit ein entscheidendes Element. Ohne Bismarck wäre sie wie „Hamlet“ ohne den Prinzen.

2. Der Disaggregierte Regulierungsansatz 1837

Das Jahr 1835 markiert die erste offizielle maschinelle Eisenbahnfahrt in Deutschland. Im Vergleich zu England, wo die Strecke Stockton-Darlington schon 1825 eröffnet worden war, und im Vergleich zu den USA, wo die erste öffentliche Eisenbahn 1826 ihren Betrieb aufnahm, war Deutschland damit zwar um ein Jahrzehnt hinterher, aber immer noch eine der Pioniernationen. In jedem Fall war die Technologie noch sehr jung und die Erfahrungen mit ihr gering. Umso verblüffender ist die Lektüre einer Schrift wie „Die Eisenbahnen und deren Actionäre in ihrem Verhältniß zum Staat“ aus dem Jahr 1837, geschrieben von David Hansemann, einem

damals führenden Politiker, Bankier und Industriellen³, mit ihren überraschend modernen ordnungspolitischen Ideen.

Im Zentrum dieser Abhandlung steht die Frage nach der optimalen Aufgabenteilung zwischen Privatwirtschaft und Staat bei Planung, Bau, Finanzierung und Betrieb der Eisenbahnen.

Hansemanns Antwort ist erstaunlich aktuell. Er zieht bereits in dieser frühen Phase des Eisenbahnbaus eine klare Trennlinie zwischen Netz und Betrieb und will die Rolle des Staates ausschließlich auf die Planung, Errichtung, Finanzierung und Instandhaltung des Netzes beschränkt wissen, ähnlich wie bei Straßen und Kanälen. Der Betrieb von Zügen auf dem Netz hingegen sei Sache der Privatwirtschaft. Hansemann schreibt:

„Man mißverstehe mich nicht. Es ist keineswegs meine Absicht, daß der Staat den Betrieb *auf* den Eisenbahnen⁴ übernehme. Dieser Betrieb

3 Es gibt vielfältige Quellen zum Leben David Hansemanns. Empfohlen sei beispielsweise der Artikel von Ott (1963), der den vorliegenden Beitrag in vielfältiger Weise vertieft. Es möge hier genügen, noch einmal stichwortartig einige zentrale Stationen seines Lebens aufzurufen. Hansemann wurde 1790 in einfachen Verhältnissen in Finkenwerder geboren und zeigte schon früh kaufmännisches Geschick. 1817 gründete er ein Textilunternehmen in Aachen, diversifizierte aber schon sehr bald in das Versicherungswesen. 1824/25 gründete er die Aachener Feuer-Versicherungsgesellschaft, die heute nach verschiedenen Umbenennungen ein Teil von AMB Generali ist. Um diese Zeit war er bereits eine bekannte und angesehene Unternehmerpersönlichkeit. In den 30er Jahren erwachte sein Interesse für Eisenbahnen und er begann sich intensiv mit dem Eisenbahnsystem anderer Länder, vor allem der USA und Belgiens auseinanderzusetzen. Ein Ergebnis dieser Studien war die bereits erwähnte Schrift von 1837. Schon bald begann er aber auch praktisch als Eisenbahnmanager und -politiker tätig zu werden, so in der Rheinischen Eisenbahngesellschaft, in der er den Posten des Vizepräsidenten von 1837 bis 1844 bekleidete. 1843 wurde er Mitglied des Rheinischer Provinziallandtags und 1847 Mitglied des Preußischen Vereinigten Landtages. Er galt als führender Vertreter des (rheinischen) Liberalismus in Deutschland, war kurz Finanzminister und von 1848 bis 1851 Präsident der Preußischen Zentralbank, wo er interessanterweise deren Privatisierung und die Herausgabe privater Banknoten befürwortete. Nach seiner Präsidentschaft widmete er sich wieder vermehrt privaten Geschäftsinteressen. 1851 gründete er die „Direction der Disconto-Gesellschaft“ (später nur „Disconto-Gesellschaft“), die sich zu einer der ersten deutschen Großbanken entwickelte und 1928 mit der Deutschen Bank fusionierte. Hansemann starb 1864. David Hansemann ist nicht zu verwechseln mit seinem Sohn, dem Bankier Adolph von Hansemann, der in der Bismarck Ära eine zentrale Rolle spielte und im Gegensatz zu seinem Vater politisch kein Liberaler war.

4 Wenn Hansemann hier von den „Eisenbahnen“ schreibt, so meint er, wie damals häufig, nicht ein bestimmtes Eisenbahnunternehmen, sondern die Schieneninfrastruktur.

ist einem großen Fabrikbetrieb nicht unähnlich, in welchem kaufmännische und technische Kenntnisse erforderlich sind. Ein Betrieb dieser Art, um vorteilhaft zu sein, muss durch Privatinteresse wesentlich unterstützt werden, auch muss in der Leitung das Maß an Freiheit der Bewegung vorhanden sein, welches einer Hoheitsverwaltung nie gegeben werden kann.“ (Hansemann, 1837, § 54, Hervorhebung vom Verfasser).

Hansemann befürwortet also eindeutig, dass der Betrieb von Zügen *auf dem Netz* privaten Unternehmen überlassen werden solle. Das Netz selbst hingegen solle entweder durch den Staat selbst oder durch private Konzessionäre bereitgestellt werden. Hansemann unterzieht beide Optionen einer näheren Analyse (Abschnitte 3 und 4 seiner Abhandlung). Zwar liegen seine Sympathien eindeutig bei der ersten Alternative, er berücksichtigt aber, dass in bestimmten Situationen budgetäre Engpässe oder andere politische Notwendigkeiten nur die zweite Lösung offen lassen, so wie es in Preußen in der Folgezeit auch tatsächlich der Fall war.

Wenn aber das Netz alleiniger Aufgabenbereich des Staates sein soll, der Betrieb hingegen Aufgabe von privaten Unternehmen, stellt sich sofort die Frage der Netzzugangsbedingungen, insbesondere die Frage nach den Trassenpreisen („Bahngeld“, wie der Ausdruck damals lautete). Hier vertritt Hansemann die gleiche Position wie etwa Hotelling sehr viel später in der Grenzkostendebatte der 1930er Jahre und plädiert für das Grenzkostenprinzip. Er argumentiert, dass die volkswirtschaftlichen Effekte der Eisenbahnen dann am größten seien, wenn die Trassenpreise entweder gleich Null gesetzt würden oder allenfalls gleich dem für Wartung und Instandhaltung entsprechenden Niveau. Seine Begründung konnte natürlich zu dieser Zeit noch nicht wohlfahrtstheoretischer Natur sein. Bekanntermaßen führte Dupuit das Konzept der Konsumentenrente erst 1844 in die Volkswirtschaftslehre ein. Stattdessen bringt er ein Argument vor, dass sich auf die positiven externen Effekte der Eisenbahntechnologie bezieht und das Anklänge an die etwa zur gleichen Zeit von Friedrich List vorgebrachte Theorie der „produktiven Kräfte“ aufweist.

Hansemann argumentiert, dass Investitionen in die Eisenbahninfrastruktur zwei Arten von volkswirtschaftlichem Kapital hervorbringen würden: Das „erste Kapital“ sei die physische Infrastruktur selbst. Das „zweite Kapital“ bestehe in den Erträgen (den „indirekten Revenuen“), die das erste Kapital erwirtschafte. In Hansemanns eigenen Worten:

„Die Zunahme des Nationalvermögens, welche indirekt unausbleiblich die Folge der Benutzung von Eisenbahnen, nämlich der verbesserten Kommunikationsmittel, sein muss, ist ein zum neu geschaffenen

Eisenbahnkapital erworbenes zweites Kapital, das erfahrungsgemäß mehr als jenes selbst beträgt. Die Staaten haben Kunststraßen und Kanäle gebaut, um das vorbezeichnete zweite Kapital zu erwerben.“ (Hansemann, 1837, § 52)

Damit das zweite Kapital hinreichend erschlossen werden könne, sei für das Bahngeld das Grenzkostenprinzip anzuwenden. Dies erfordere aber aufgrund des entstehenden Defizits den Bau und Betrieb durch den Staat.

Bekanntermaßen reicht dieses Argument für sich alleine nicht aus, um ein Staatsmonopol zu begründen. Auch private Netzanbieter könnten das Grenzkostenprinzip anwenden. Das entstehende Defizit – sofern es überhaupt auftritt⁵ – könnte aus der Staatskasse beglichen werden. Hansemann lässt diese Möglichkeit außer Acht, möglicherweise aus Gründen der engen budgetären Handlungsspielräume in seiner Zeit.

Ein Staatsmonopol beim Netz war nach Hansemann aber auch deshalb wünschenswert, weil es die Quersubventionierung unprofitabler Strecken durch profitable zuließ. Er fürchtete, wie auch List zur gleichen Zeit, dass private Netzbetreiber nicht bereit sein würden, neben „dichten“ auch „dünne“ Verbindungen zu errichten. Die spätere Entwicklung zeigte, dass diese Befürchtungen nicht ganz unbegründet waren. Beispielsweise waren private Unternehmen zunächst nicht bereit, die dünnbesiedelten Gebiete Ostpreußens an das bereits vorhandene Eisenbahnnetz anzubinden. (Die Frage der „Ostbahn“.) Solchen Fehlentwicklungen, so Hansemann, könne der Staat durch ein Netzmonopol entgegenwirken, da er in Abwesenheit von Wettbewerb die Verluste bei den dünnen Strecken durch Gewinne bei den dichten Strecken ausgleichen könne (Hansemann, 1837, § 47).

Es ist nicht ganz klar, wie Hansemann diese Quersubventionierung mit dem von ihm befürworteten Grenzkostenprinzip bei den Trassenpreisen vereinbaren will. Möglicherweise denkt er hier eher an eine Subvention verlustbringender Netzteile aus einer Besteuerung auf der Betreiberebene. Eine solche Besteuerung wäre aber auch unter Netzwettbewerb möglich, etwa unter Einsatz des von Knieps befürworteten Universaldienstfonds.

Hansemann war mit seiner regulatorischen Differenzierung zwischen Netz und Betreiberebene zum damaligen Zeitpunkt zweifellos sehr fortschrittlich, aber keineswegs ein einsamer Visionär. Er verweist selbst in seinem Vorwort darauf hin, dass er sich viele seiner Ideen aus dem einschlägigen Schrifttum und aus Gesprächen mit zuständigen Beamten angeeignet

5 Im mikroökonomischen Firmenoptimum wird bekanntlich die Kapazität so gewählt, dass sich langfristige und kurzfristige Grenz- und Durchschnittskostenkurven schneiden.

habe. Wie eingangs schon erwähnt, findet sich auch im Preußischen Eisenbahngesetz von 1838 die Idee eines Wettbewerbs auf dem Netz. § 27 des Gesetzes bestimmte, dass nach einer Wartezeit von 3 Jahren auch andere Unternehmen als der Netzerbauer Zugang zu einem bereits errichteten Netz haben konnten. Voraussetzung war der Erwerb einer entsprechenden Konzession vom zuständigen Ministerium und die Entrichtung einer Netzzugangsgebühr („Bahngeld“) an den Netzbetreiber. Die genaue Berechnungsmethode des Bahngeldes war im Gesetz in den folgenden §§ 29 und 30 detailliert spezifiziert (Fremdling & Knieps, 1993, S. 133, 144).

Doch zurück zu Hansemanns Abhandlung. Nachdem er ein Monopol auf der Netzebene begründet hatte, mit gleichzeitigem Wettbewerb auf der Betreiberebene, stellte sich als nächstes die Frage nach der Ausgestaltung dieses Wettbewerbs. An dieser Stelle brems nun Hanseman seinen Reformeifer etwas. Statt offenem Wettbewerb (heute: „open access“) befürwortet er beim Betrieb von Zügen ein Gebietsmonopol. Seine Gründe sind einerseits Sicherheitsbedenken, andererseits ökonomisch-technologische Synergien (Verbundvorteile) zwischen Netz und Betrieb. Er zitiert zeitgenössische empirische Befunde aus anderen Ländern, vor allem Pennsylvanien, die zu belegen schienen, dass offener Wettbewerb zu einer höheren Zahl an Unfällen führte als ein Betreibermonopol (Hanseman, 1837, § 87). Hinsichtlich der Synergien weist er darauf hin, dass die Instandhaltungskosten des Netzes entscheidend von den technischen Eigenschaften und der Qualität des sich darauf bewegenden Rollmaterials abhängen. Eine Kostenoptimierung für das gesamte „Rad-Schiene-System“ (wie es ein Jahrhundert später heißen sollte) sei nur unter einem Betreibermonopol möglich. Beide Argumente sind in unserer Zeit umfassend erörtert worden, so auch von Knieps, und sollen deshalb hier nicht weiter vertieft werden.

Nach welchen Kriterien soll nun aber der monopolistische Betreiber ermittelt werden? Welche Prinzipien sollen für die „Konzessionierung des Betriebs *auf* Eisenbahnen“ zum Einsatz kommen (Hanseman, 1837, § 56, Hervorhebung vom Verfasser). Hier überrascht Hanseman erneut, indem er eine Art von Demsetz-Auktion vorschlägt, mehr als 150 Jahre bevor Demsetz diese Idee in die volkswirtschaftliche Regulierungstheorie einführte.

Zunächst stellt er die Anforderungen auf, die in der zu versteigernden Konzession festzuschreiben seien (Hanseman, 1837, § 56).

Die Lizenz müsse, erstens, eine Maximaldauer enthalten, die aber lang genug sein müsse, um dem Betreiber hinreichende Investitionsanreize zu geben, damit „zum allgemeinen Wohle in der Folge alle Vervollkommnungen des Betriebs dem Publikum zu Gute kommen“ (Hanseman,

1837, § 56). Hansemann sieht hier eine Frist von 20 Jahren als ausreichend an. Ferner müsse sich der Lizenznehmer verpflichten, die Eisenbahn in „vollkommen gutem Zustande“ zu erhalten.⁶ Weiterhin müsse der Betreiber zusagen, alle Betriebsanlagen an seinen Nachfolger zu einem festgelegten Preis zu übergeben, dessen Berechnungsmethode im Voraus festzulegen sei. Hansemann spricht hier also explizit schon klar die Frage der Markzutrittsbarrieren bzw. der Sunk Costs bei Auktionen an, die auch im Ansatz der disaggregierten Regulierung eine zentrale Rolle spielen.

Diese Konzession sei nun in einem Verfahren zu auktionieren, bei dem der Staat zwischen drei Zielen wählen müsse.

Der Staat könne, erstens, das Ziel haben, die Auktionserlöse zu maximieren. In diesem Falle solle die Lizenz jenem Betreiber erteilt werden, der das höchste Angebot mache.

Der Staat könne, zweitens, das Ziel haben, eine bestimmte Mindestrendite auf das eingesetzte Kapital zu erwirtschaften. In diesem Fall solle die Lizenz an denjenigen Bieter gehen, der diese Rendite garantiere, gleichzeitig aber die niedrigsten Transporttarife anbiete.

Der Staat könne schließlich, drittens, möglichst niedrige Transporttarife anstreben. Dann sei die Lizenz jenem Betreiber zu erteilen, der diesbezüglich das niedrigste Angebot macht. Dieser dritte Ansatz entspricht vermutlich am ehesten der Lehrbuchdarstellung einer Demsetz-Auktion. Allerdings hat Hansemann bei den Eisenbahntarifen nicht einen uniformen Tarif im Auge, wie in der klassischen Demsetz-Auktion. Er plädiert vielmehr für differenzierte Tarife, insbesondere im Sinne der traditionellen Wertstaffel. Dies macht die Auswahl des besten Angebots natürlich sehr viel schwieriger, wozu er sich allerdings nicht weiter äußert.

Dieser Vorschlag Hansemanns ähnelt in der Tat sehr stark einer Demsetz Auktion. Der Verfasser hat deshalb vorgeschlagen, Hansemann als Vorläufer von Demsetz anzusehen und ihm sogar Priorität gegenüber Chadwick (1800–1890) zuzusprechen, der in einem Artikel von 1859 ähnliche Ideen entwickelt hat und erstmals den Unterschied zwischen dem Wettbewerb im Markt („competition within the field“) und dem Wettbewerb um einen Markt („competition for the field“) geprägt hat.⁷ Dem steht jedoch entgegen, dass z.B. Friedrich List in seinen Eisenbahnschriften erwähnt, dass solche Auktionen zu dieser Zeit schon routinemäßig in anderen Ländern, so etwa in Frankreich, angewandt wurden, nicht nur im Eisenbahnwesen sondern auch in anderen Bereichen. Die Prioritätsfra-

6 „Eisenbahn“ bezieht sich hier wieder auf das Schienennetz, s. Fn. 4.

7 Auf die Priorität von Chadwick hat als erstes Coase hingewiesen, s. Kitch (1983).

ge scheint also hier, wie so häufig in der Wissenschaft, kaum mehr zu entscheiden zu sein. Dies ändert nichts an Hansemanns Verdiensten.

3. Verstaatlichung statt disaggregierter Regulierung: Die Rolle Bismarcks

Hansemann konnte sich weder mit seiner Idee einer ausschließlichen Bereitstellung der Schieneninfrastruktur durch den Staat, noch mit seinem Trennmodell, noch mit seiner Auktionsidee durchsetzen. Die Entwicklung verlief in der Folgezeit gänzlich anders.

Vor die politische Schwierigkeit gestellt, aufgrund des Staatsschuldengesetzes von 1820 keine Anleihen begeben zu können, setzte die preußische Regierung zunächst ausschließlich auf privates Kapital. Konzessionen gingen allein an private vertikal integrierte Unternehmen, die Transportleistungen von der Infrastruktur bis zum Betrieb aus einer Hand anboten (so übrigens auch an die Rheinische Eisenbahn, an deren Spitze Hansemann von 1837 bis 1844 stand). Um die entsprechenden Investitionen für die Aktionäre attraktiv zu machen, war im Preußischen Eisenbahngesetz von 1838 vorgesehen, dass für einen Zeitraum von 30 Jahren keine parallel laufenden Eisenbahnlinien konzessioniert würden. Konsequenterweise wurde auch Wettbewerb auf der Betreiberebene nicht zugelassen, obwohl, wie erwähnt, das Gesetz die Möglichkeit dazu enthielt. Wettbewerb auf dem Netz hätte zu kostenorientierten Tarifen auf der Betreiberebene geführt und damit die Investitionsanreize der Privatwirtschaft verringert.

Ab 1850 traten auch staatliche Unternehmen bzw. quasi-staatliche Unternehmen in den Markt ein, bzw. entstanden durch Kauf von in finanzielle Schieflage geratenen Bahnen. Besondere Erwähnung verdient hier die Bergisch-Märkische Eisenbahn, die ebenfalls vom Staat übernommen wurde und, obwohl von Staatsbeamten geleitet, sich als besonders aggressiver und erfolgreicher Wettbewerber erwies.⁸

Zu dieser Zeit waren allerdings die Restriktionen hinsichtlich parallel laufender Strecken schon deutlich gelockert, und es kam zu einer Fülle von neuen Marktzutritten. Ziel war es, die rasant wachsende Nachfrage nach Eisenbahntransporten zu befriedigen, gleichzeitig aber den Missbrauch etwaiger Monopolstellungen zu beschränken (Fremdling &

8 Ähnliche Effizienzwirkungen des aktuellen und potentiellen Wettbewerbs auf öffentliche Unternehmen wurden im letzten Jahrhundert auch in anderen Branchen beobachtet, so im Luftverkehr (etwa die Lufthansa und andere Flag-Carrier) und in der Telekommunikation (Televerket in Schweden).

Knieps, 1993, S. 131, 133). Mit Beginn der 1870er Jahre hatte diese Entwicklung jedoch zu Überkapazitäten und ruinöser Konkurrenz geführt, in der mehrere Unternehmen in erhebliche wirtschaftliche Probleme gerieten. Begleiterscheinung war ein Tarifchaos, in dem teilweise über 1500 verschiedene Tarife nebeneinander existierten, wodurch die Durchführung überregionaler Transporte stark behindert wurde (Fremdling & Knieps, 1993, S. 140). Von der verladenden Wirtschaft wurde immer stärker eine Vereinheitlichung gefordert. Gleichzeitig waren Kollusion und Kartellierungen die Folge der schwierigen Gesamtsituation. Eine grundlegende Reform schien unvermeidlich.

An dieser Stelle wird nun ein Blick über die Eisenbahnpolitik hinaus auf die politische Großwetterlage unumgänglich. Der Verfasser ist hier etwas zögerlich, den Schritt in fachfremdes Territorium zu wagen, einige skizzenhafte Bemerkungen mögen jedoch vertretbar sein. Quelle ist in erster Linie Fritz Sterns monumentale Doppelbiographie Bismarcks und seines Bankiers Bleichröder (s. Stern, 2000).⁹

In der Zeit, über die wir hier reden, hatte es in Preußen dramatische politische Veränderungen gegeben. Bismarck war an die Macht gekommen und Preußen hatte drei Kriege geführt, in denen die militärische Bedeutung der Eisenbahnen klar zum Ausdruck gekommen war. Außerdem hatte sich die wirtschaftspolitische Grundstimmung radikal verändert. Um 1868 hatte zunächst ein Aufschwung eingesetzt, der nach Gründung des Deutschen Reiches 1871 in eine extreme Boomphase einmündete und, wie häufig in solchen Zeiten, mit zahlreichen negativen Begleiterscheinungen verbunden war. „Es war eine Zeit echter Expansion, betrügerischer Gründungen, eine Ära halb legaler Verwegenheit, als der Reiz des Profits normale Besonnenheit ausschaltete. Bald tauchte das Schlagwort der Gründerzeit auf, „eine Epoche, in der neue Imperien in Politik und Wirtschaft gegründet wurden“ (Stern, 2000, S. 265). Auch der Eisenbahnbereich blieb von fragwürdigen Spekulationen nicht verschont. Am 14. Januar 1873 legte der Abgeordnete Eduard Lasker, ein Intimfeind Bismarcks, in einer Parlamentsrede zahlreiche Skandale bei der Vergabe von Eisenbahnkonzessionen offen. Die Konsequenzen dieser Enthüllungen waren, erstens, der Rücktritt des zuständigen Handelsministers und, zweitens, die Einsetzung einer Untersuchungskommission, deren Ergebnisse in die Empfehlung einer Verstaatlichung der Eisenbahnen mündeten.

9 Es gibt selbstverständlich ein spezialisiertes Schrifttum zur Eisenbahnpolitik Bismarcks, das aber eher historischer als ökonomischer Natur ist. Der Klassiker ist von der Leyen (1914).

Die Höhenflüge der Gründerzeit endeten jäh mit dem „Wiener Börsenkrach“ von 1873, der mehrere weitere Börsenkräche nach sich zog, so auch in Berlin. Eisenbahnaktien gehörten zu denjenigen Papieren, deren Kurse am stärksten in Mitleidenschaft gezogen wurden. „Anfang 1874 gaben 61 Banken, 116 industrielle Unternehmen und vier Eisenbahngesellschaften ihren Bankrott bekannt.“ (Stern, 2000, S. 266) Es folgte eine der längsten weltweiten Depressionsphasen der neueren Wirtschaftsgeschichte.

Die Wirtschaftskrise erschwerte einerseits die Finanzierung der teilweise hochverschuldeten privaten Eisenbahnen, andererseits begannen spätestens mit dieser Entwicklung die liberalen wirtschaftspolitischen Ideen an Einfluss zu verlieren, die noch in den 1850er und 1860er die Wirtschaftspolitik und damit auch die Eisenbahnpolitik geleitet hatten. Die führenden Vertreter verloren nach und nach Macht und Ämter.

Bismarck selbst hatte niemals große Sympathien für marktwirtschaftliche Prinzipien gehabt. Seine wirtschaftspolitischen Vorstellungen – sofern er überhaupt durchdachte Vorstellungen in dieser Richtung hatte – waren protektionistischer, interventionistischer und paternalistischer Natur. Im Fall der Eisenbahnen ging es ihm um einheitliche Tarife, einheitlichen Betrieb und wirksamere Nutzung im Kriegsfall. Dementsprechend begann er einen langen und zähen Kampf um die Verwirklichung eines Staatsbahnsystems im ganzen Deutschen Reich.

1873 wurde vom Parlament auf sein Betreiben hin das Kaiserliche Reichseisenbahnamt geschaffen. Vordergründig war diese Gründung eine Reaktion auf die Rede Laskers und auf den öffentlichen Druck in Richtung einer Vereinheitlichung der Tarife. Die tiefere Absicht Bismarcks dürfte aber die Vorbereitung eines Staatsbahnsystems gewesen sein. Er traf dabei allerdings zuerst auf erbitterten Widerstand. Das Parlament und die Bundesstaaten setzten sich massiv dagegen zur Wehr, dem Amt weiterreichende Befugnisse zuzugestehen. Der Versuch, ein Reichseisenbahngesetz auf den Weg zu bringen, wurde abgeblockt.

Erst 1876 gelang es Bismarck, dem Landtag ein Gesetz abzurufen, das den Verkauf der preußischen Eisenbahnen an das Deutsche Reich vorsah. Es erhob sich allerdings erregter Protest, vor allem aus liberalen Kreisen. Ein Kommentator schrieb, dass der Kanzler „mit dem Project dem Socialismus und Communismus unbewußt in die Hände“ arbeite. „Es ist ein mit kommunistischer Kraftbrühe gekochtes Gericht“ (Stern, 2000, S. 303).

Trotz seiner mannigfaltigen Bemühungen und sogar einer zwischenzeitlich ausgesprochenen Rücktrittsdrohung, gelang es Bismarck nicht, seine Verkaufspläne an das Deutsche Reich zum Erfolg zu führen. Die legendäre Durchsetzungskraft des eisernen Kanzlers fand ihre Grenzen in den Widerständen, die ihm das Parlament, die eigenen Kabinettskollegen

und die Bundesstaaten entgegensetzten. Er reduzierte deshalb 1878 seine Politik auf die Strategie, *wenigstens in Preußen* die in der Wirtschaftskrise verbliebenen Privatbahnen aufzukaufen. Grundgedanke war es, auf diesem Wege zumindest dort ein monolithisches staatliches Eisenbahnsystem zu schaffen, das durch seine schiere Größe und Dominanz auch den anderen deutschen Bundesstaaten den Weg zur Verstaatlichung weisen würde.

Es erhebt sich die Frage, welche Argumente oder welche Einflüsse Bismarck zu seinem Kreuzzug für eine Verstaatlichung der Eisenbahnen letztlich bewogen. Fritz Stern behauptet, dass die tatsächliche Geschichte seiner Eisenbahnpolitik „noch nicht umfassend aus der Sicht der Gegenwart“ behandelt worden sei. (Stern, 2000, S. 302). Der Verfasser ist kein hinreichender Bismarck Kenner, um sich hier ein Urteil erlauben zu können. Dass der militärische Aspekt eine entscheidende Rolle spielte, ist offensichtlich und wurde bereits erwähnt. Inwieweit ihn wirtschaftstheoretische Überlegungen beeinflusst haben, ist weniger klar. In seiner Rede vor dem Preußischen Herrenhaus am 18. Mai 1876 erwähnt er ein Argument, das in moderner Sprechweise als Argument des natürlichen Monopols charakterisiert werden könnte. Er sieht die in vielen Ländern zu beobachtende zurückgehende Rentabilität der Bahnen als Folge, „des Irrthums, dass die Concurrenz in Eisenbahnlinien den Verkehr wohlfeiler macht.“ Bei gegebenem Verkehrsvolumen, müsse eine Duplizierung von Verbindungen zwangsläufig dazu führen, dass die Kosten pro transportierter Einheit steigen. Auf diese Weise, so Bismarck, habe die Entwicklung dazu geführt, dass die Aktionäre, „zum großen Theil froh sein“ würden, „eine zurückgehende Rente zu einem festen Course, wenn der Preis annehmbar ist, loszuschlagen.“ Widerstand sei nur von den Eisenbahndirektoren zu befürchten, die den Verlust ihrer Positionen und ihrer Einkommen zu vermeiden wünschten.

In einer früheren Rede vom 26. April des gleichen Jahres vergleicht Bismarck den Zustand des Eisenbahnwesens mit seiner Vielfalt von Zuständigkeiten und Tarifen mit dem Mittelalter. Die Zersplitterung des Eisenbahnwesens münde in einen ökonomisch schädlichen Machtkampf der Eisenbahndirektionen untereinander, der zu beenden sei. In der gleichen Rede klingen auch gemeinwirtschaftliche Motive an. Die Eisenbahnen seien nicht dazu da, in erster Linie „ein Gegenstand finanzieller Concurrenz zu sein, um das Höchstmögliche herauszuschlagen“; sie seien „viel mehr für den Dienst des Verkehrs, als für den Dienst der Finanzen bestimmt.“¹⁰ Es

10 Beide Reden werden zitiert nach Kohl (1893, S. 383 ff., 408 ff.).

muß weiterer Forschung vorbehalten bleiben, die intellektuellen Wurzeln der bismarckschen Eisenbahnpolitik genauer offenzulegen.

Ungeachtet mannigfaltiger politischer und praktischer Schwierigkeiten bei seiner Umsetzung war der Aufkaufsplan Plan des Kanzlers letztlich erfolgreich. „Bismarcks Plan glückte: die Eisenbahnen, ein Lebensnerv der nationalen Wirtschaft im Frieden und der Verteidigung im Krieg mußten dem preußischen Staat gehören. 1876 waren 4683 km Staatseigentum, um 1890 hatte sich das Eisenbahnnetz durch Ankäufe für mehr als 2,8 Milliarden Mark auf 14 000 km erweitert. Private Linien gab es kaum mehr. Der Staat unterhielt ein vorbildliches Verkehrssystem, das als leistungsfähig, zuverlässig und wirtschaftlich bekannt war. ... Alles in allem hatte Bismarck Grund zur Zufriedenheit mit seinem Erfolg ...“ (Stern, 1988, S. 312).

Grund zu Zufriedenheit hatte Bismarck auch rein persönlich. In Befolgung der amerikanischen Weisheit „put your money where your mouth is“ hatte er einen erheblichen Teil seines Vermögens in Eisenbahnpapieren angelegt. Ebenso sein persönlicher finanzieller Berater, der Großbankier Gerson Bleichröder. „Drei Tage nach Inkrafttreten des Gesetzes, das Preußen autorisierte, seine Eisenbahnlinien dem Reich zu verkaufen, instruierte Bismarck Bleichröder, für 30 000 Mark 4,5 %ige Vorzugsaktien der Berlin-Stettiner Eisenbahn und um den gleichen Betrag 4,25 %ige der Berlin-Anhalter Eisenbahn zu kaufen. Es ergab sich, daß die Strecke Berlin-Stettin einige Jahre später als die erste in Preußen verstaatlicht wurde. Im Lauf von acht Jahren kaufte und verkaufte Bismarck ... nacheinander Eisenbahnaktien für mehr als eine Million Mark; manchmal war etwa die Hälfte seines flüssigen Kapitals in solche Aktien investiert – die klarste Festlegung auf seine Politik der Verstaatlichung; ein Fehlschlag dieser Politik oder eine übermäßige Verzögerung der Verstaatlichung hätte ihn Geld kosten können“ (Stern, 2000, S. 305). Bleichröder seinerseits kaufte von 1875 an „so diskret wie möglich“ (Stern, 2000, S. 304) Aktien von Gesellschaften auf, die vermutlich verstaatlicht würden. Die praktische Durchführung der entsprechenden Börsenmanöver vertraute er seinem damaligen Mitarbeiter Carl Fürstenberg an, der später zu einem der führenden Großbankiers des Deutschen Reiches werden sollte und der sich hier seine ersten Sporen verdiente. (Für weitere Transaktionen siehe ausführlich Stern, 2000) Diese Geschäfte Bleichröders waren allerdings niemals ganz ohne Risiko. Die meisten der Bahnen steckten in finanziellen Schwierigkeiten und es war nie ganz sicher, ob der Staat sie auch wirklich zu einem guten Preis aufkaufen würde. Für viele Bahnen war der Aufkauf durch den Staat die letzte Überlebenschance.

Handelte es sich hier um Insiderhandel? Stern weist darauf hin, dass die Verstaatlichungspläne allgemein bekannt waren. Jeder interessierte Anleger konnte im Prinzip ebenso gut in Aufkaufskandidaten investieren, wie der Reichskanzler. Unklar war aber, wann und zu welchen Konditionen der Kauf einer bestimmten Eisenbahn stattfinden würde. „Darüber wußte Bismarck besser Bescheid als sonst jemand; außerdem hatte er die Macht, seine Minister unter Druck zu setzen“ (Stern, 2000, S. 311). Stern will in den Investments Bismarcks dennoch keinen Insiderhandel erblicken, da der Kanzler ungeachtet seiner Machtfülle und seines Informationsvorsprunges niemals sicher sein konnte, mit seinen Plänen auch durchzudringen. Jeder Einzelfall war vom Parlament zu genehmigen. Bismarck habe insofern nicht besser gestanden als jeder andere Anleger. Stern gesteht allerdings zu, dass nach heutigen Begriffen Bismarcks Aktienkäufe vermutlich eher als problematisch einzuschätzen seien. Der Leser möge sich hier sein eigenes Urteil bilden.

Als gesichert kann gelten, dass Bleichröder von seinem engen Vertrauensverhältnis zu Bismarck sehr wohl zu profitieren wusste (Einzelheiten s. Stern, 2000). Auch seine vornehmsten Kunden, waren sich über die Sonderstellung des Bankiers im Klaren und ersuchten um vertrauliche Informationen. Prinz Heinrich VII von Reuß etwa fragte im Mai 1879 nach, ob er nicht Köln-Mindener Aktien zum Kurs von 118 abstoßen solle, nachdem Bleichröder geäußert habe, der Kurs werde bei einer Verstaatlichung auf etwa 115 zu stehen kommen. Heinrich Graf von Lehndorff berichtete Bleichröder von einem Gespräch mit dem für die Eisenbahnen zuständigen Minister für öffentliche Arbeiten, Maybach, in dem er vertraulichen Informationen über die bevorstehende Verstaatlichung der Rheinischen und Anhalter Linien bekommen hatte. „Meinen Sie nun, daß es lohnte, an den Kauf Rheinischer Aktien zu denken?“ Bleichröder seinerseits hielt Bismarck über die Nöte seiner Kunden auf dem Laufenden und regte gegebenenfalls auch die Beseitigung personeller Widerstände an. So wies er Bismarck in einem Brief vom Juni 1879 darauf hin, dass der Finanzminister Hobrecht im vertraulichen Gespräch mit Minister Maybach bekundet habe, weiteren Aufkäufen nicht mehr zustimmen zu wollen. „Einen Monat danach war Hobrecht nicht mehr im Amt, der entgegenkommendere Karl Bitter übernahm das Finanzministerium“ (Stern, 2000, S. 307).

Umtriebe dieser Art gab und gibt es natürlich zu allen Zeiten und auf allen Gebieten. Die preußische Eisenbahnpolitik am Ende des 19. Jahrhunderts stellt hier keinen Sonderfall dar. Worauf es an dieser Stelle ankommt, ist vielmehr, dass diese Politik mit Sicherheit keinen Zusammenhang zur akademischen regulierungstheoretischen Diskussion mehr hatte. Ihre zen-

trale Antriebskraft lag in den machtpolitischen und persönlichen Interessen Bismarcks und der Führungselite des preußischen Staates.

4. Schlussbemerkungen

Milton Friedman hat die Rolle akademischer Ökonomen einmal mit folgenden Worten charakterisiert:

„That, I believe, is our basic function: to develop alternatives to existing policies, to keep them alive and available until the politically impossible becomes politically inevitable” (Friedman, 2002, XI).

Die beiden vorangegangenen Abschnitte scheinen ihm Recht zu geben. Hansemann konnte mit seinen Vorstellungen in den 1830er Jahren im damaligen politischen Umfeld nicht durchdringen, so gut durchdacht sie auch immer gewesen sein mögen. Das Gleiche gilt für die Eisenbahnreformer der 1870er Jahre, deren Scheitern Fremdling und Knieps so beredt beklagen. Ideen dieser Art trafen erst in der zweiten Hälfte des 20. Jahrhunderts auf ein günstiges Klima, als zumindest in Europa das Staatsbahnsystem immer stärker in die Kritik geriet und die Zeichen nicht nur im Eisenbahnwesen sondern in vielen anderen Wirtschaftszweigen in Richtung Liberalisierung und Deregulierung standen. Im politischen Klima der Bismarck-Zeit war daran nicht zu denken.

Noch eine weitere Lehre scheint die Preußische Eisenbahngeschichte bereit zu halten. Auch Wirtschaftsgeschichte bleibt zuletzt immer Teil der Gesamtgeschichte. Ohne ein Verständnis der staatspolitischen Ziele Bismarcks und der ihm nahestehenden Kreise ist die preußische Eisenbahnpolitik unverständlich. Dass er von seiner Politik schließlich auch persönlich profitieren konnte, bleibt ein Nebeneffekt. Um noch einmal Stern zu zitieren: „... wirtschaftliche Erfordernisse, politisches Kalkül und persönliche Vorlieben taten sich zusammen, seinen Kurs in den späten 1870er Jahren und für das kommende Jahrzehnt zu bestimmen“ (Stern, 2000, S. 301).

Die zentrale These von Fremdling und Knieps wird durch die Einbeziehung der Rolle Bismarcks nicht in Frage gestellt, sondern eher gestärkt. Die Geschichte des preußischen Eisenbahnwesens taugt nicht als Beleg gegen Wettbewerb im Eisenbahnwesen. Zwar ist richtig, dass der Wettbewerb vertikal integrierter Eisenbahnunternehmen schließlich zu Problemen führte, andere wettbewerblich orientierte Reformansätze im Geist von Hansemann oder vom Typ des disaggregierten Regulierungsansatzes wurden aber nie versucht. Dies hing nicht mit inhärenten Problemen dieser Ansätze zusammen. Sie scheiterten vielmehr am wirtschaftspolitischen

Zeitgeist in der Depressionsphase der 1870er Jahre und den wirtschafts- und staatspolitischen Ideen Bismarcks.

Referenzen

- Fremdling, R., & Knieps, G. (1993). Competition, regulation and nationalization: The Prussian railway system in the nineteenth century. *Scandinavian Economic History Review*, 41(2), 129–154. <https://doi.org/10.1080/03585522.1993.10415864>
- Friedman, M. (2002). *Capitalism and Freedom* (40th Anniversary Edition). Chicago, IL & New York, NY: University of Chicago Press.
- Hansemann, D. (1837). *Die Eisenbahnen und deren Aktionäre in ihrem Verhältnis zum Staat*. Leipzig & Halle: Renger'sche Buchhandlung.
- Kitch, E. W. (1983). The Fire of Truth: A Remembrance of Law and Economics at Chicago, 1932–1970. *The Journal of Law & Economics*, 26(1), 163–234.
- Knieps, G. (2001). *Wettbewerbsökonomie*. Berlin & Heidelberg: Springer.
- Knieps, G. (2007). *Netzökonomie*. Wiesbaden: Gabler.
- Kohl, H. (Ed.) (1893). *Die politischen Reden des Fürsten Bismarck* (Sechster Band, 1873–1876). Stuttgart: Cotta Stuttgart.
- Ott, K. (1963). Die Eisenbahnen in ihrem Verhältnis zum Staat. *Archiv für Eisenbahnwesen*, 73(3), 262–290.
- Stern, F. (2000). *Gold und Eisen: Bismarck und sein Bankier Bleichröder*. Reinbek: Rowohlt.
- Von der Leyen, A. (1914). *Die Eisenbahnpolitik des Fürsten Bismarck*. Berlin & Heidelberg: Springer.
- Wieland, B. (2018). Two early views on railway regulation in Germany: Friedrich List and David Hansemann. In: H. Hagemann, S. Seiter & E. Wendler (Eds.). *The Economic Thought of Friedrich List* (pp. 122–136). London: Routledge.

Access regulation of non-personal data markets

Patrick Zenhäusern¹

Abstract

Non-personal data like financial market data or smart-farming data often are a critical production input for entrepreneurial activity. Access to such data can either be unrestricted (“open data”) or restricted. If certain non-personal data is required in the production of a particular product or service and if it cannot be duplicated with reasonable deployment of resources, competition in this market is not possible without third-party access to this data.

From an economic perspective, third-party access rights to non-personal data can be legitimised if the data has the characteristics of “essential data”. If non-personal data access is impeded or even refused through the exploitation of market power, competition authorities and civil courts in the EU and Switzerland currently are already allowed to impose ex post penalties or approve indemnity claims for ex post behaviour found to be unlawful.

Given the disadvantage of investigating only past abuses with the current authority practice, the introduction of market investigations should be examined as an additional instrument of competition law, following the competition authorities in the UK or the Netherlands. As long as the abuse of market power in the context of data access does not occur systematically, however, there is no need for sector-specific ex ante regulation, such as mandatory licensing.

Keywords:

non-personal data, market power, essential data, market investigations, mandatory licence

1 Polynomics AG, Member of the Strategy Board;
patrick.zenhausern@polynomics.ch.

1. Introduction

The creation of value in business is changing with the development of the Industry 4.0 and the Internet of Things (IoT). Such developments lead to an increasing use of technology in and between the sectors. Data plays an important role in the search for new products and services. In addition to the sheer volume of data (and the development of the technology required to use it), data access rights are also of great importance (Weber & Thouvenin, 2018; UVEK, 2019).

Data research often focuses on personal data, where availability is often limited by data protection. In this article, we focus on non-personal data, where the primary question is the access to such data. Increasing digitisation and related processes, business activities, but also governmental and private action: they all produce non-personal data such as meteorological data, topographical data, data on pedestrians and traffic flows, data produced by machines (machine tools) and industrial product chains (European Commission, 2019a, pp. 5ff.).

Non-personal data can be negatively defined as “any data that is not personal data” (IGE, 2021, p. 6). However, it cannot always be abstractly and unambiguously determined whether data is non-personal data or not. Even though anonymised and aggregated data seems to be obvious candidates to qualify as non-personal data at first glance, it has to be ensured that it remains virtually impossible to associate it again to natural persons. When analysing both personal and non-personal data together, from a regulatory point of view, the data is defined as a “mixed data set”.²

Non-personal data is often essential to ensure that associated products and services function properly, but they also play a key role to advance research or bear potential to develop other goods and services. Not only private players, but also research institutions and authorities are interested in using non-personal data. As improved access to data can have significant economic and social advantages, potential (regulatory) interventions for improving data availability are currently being discussed (see, e.g., Frontier Economics, 2021).

2 In the EU, the personal data included in mixed data sets is governed by the “General Data Protection Regulation” (European Parliament and the Council, 2018a) and the non-personal data included in data sets by the “Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union” (European Parliament and the Council, 2018b). If the data is “inseparably linked”, the General data Protection Regulation applies fully to the entire mixed data set.

This article deliberates the question of access regulation from a perspective of competition and regulatory economics. With reference to a report by Polynomics (Schneider et al., 2020) it mostly focusses on the situation in Switzerland and the European Union, but the results are generally valid. The theoretical findings are enriched with insights from expert talks with ABB, EMS Chemie, Fenaco, Lombard Odier, Migros, Schneider Electric & Feller, Postfinance, Schindler, the City of Lucerne and Swisscom. We also include key findings with regard to the relevance of non-personal data in business using results from an online survey amongst Swiss companies in different sectors.³

We first describe the markets for non-personal data (section 2) and then go on to investigate to which extent these markets could actually fail (section 3). In section 4 we examine the dynamics in the non-personal data markets as well as the importance of investment incentives. And we explain why (from the perspective of competition economics) regulated data access is only justified under special circumstances. In section 5 we then present key implications for competition and regulatory economics. Section 6 concludes.

2. Markets of non-personal data

For companies, the actual value of data results from the integral value creation of collecting, storing, processing and analysing it. The decisive factor is their idea of an application based on the data. This idea can already be known *ex ante*, i.e., before the collection of the data such that the idea determines which data is generated, stored and analysed. But it is also possible that findings from the analysis of existing data result in a new product or service. Then the idea manifests after data collection such that the available data determines what can be analysed.

Because third parties (such as another company, a research facility, the authorities or a private person) could also have valuable ideas on how to use the data in a potentially different context and possibly in connection with other data sources, a broader access to data can increase the overall economic benefit. As a consequence, the access to data is important. In

3 The survey design is based on a company survey by the European Commission on the topic of “B2B Data Sharing Principles and Guidance” (European Commission, 2019b). The part of the survey on the Internet of Things is based on the “Staff Document COM(2018) 232 final” of the European Commission (European Commission, 2018).

the following, we therefore discuss specific options for the organization of non-personal data markets.

2.1 Primary markets

Data is often produced through the use of services. Thus, data can be obtained directly from the users of the services. Markets that provide such data are called “primary markets” (Schweitzer & Peitz, 2017, p. 22). Our expert talks show that companies acquire data both automatically (i.e., through production and operations) and deliberately (i.e., through the use of sensors). The latter is necessary, for example, in the agricultural sector to provide retailers with a proof of origin for their products or with information to which extent they have been produced organically or sustainably.

The Internet of Things (IoT) continuously generates non-personal data and therefore takes on a central role in the context of primary markets. IoT devices such as bi-directional measurement devices, sensors, actuators and remote controls use interactive machine-to-machine communication, typically in real time. The online company survey that we conducted in Schneider et al. (2020) shows that in the mechanical engineering sector, more than 30 percent of Swiss companies already use IoT devices and more than 25 percent in the telecommunications and IT sector. In the other sectors, however, less than 10 percent of companies use this technology.

2.2 Secondary markets

The term “secondary market” refers to a negotiated data access by third parties. They either have generated this data themselves or have purchased it and may also use it themselves. Data can be traded directly or via intermediary platforms in the secondary markets. The data transfer between data providers and buyers is simplified through technical functions and standardised interfaces (Schweitzer & Peitz, 2017, p. 22). This procedure is particularly advantageous for public and sector-specific data. Examples include weather data or financial market data.

In the agricultural sector, third parties typically purchase weather data and forecast models for the harvest. Weather and soil data can be primarily used for determining how to optimally fertilise fields and which produce

gives the best yield. Based on forecasts, agricultural businesses can sell their harvest produce prior to the harvest through forward transactions.

Secondary markets are also important in the financial sector. Data aggregators such as SIX, Bloomberg or Reuters use global raw data of stock exchanges and other exchanges, aggregate it into data feeds and thus form products which they sell to financial service providers. The fees for such services has decreased significantly in recent years according to experts. As compensation for the decreasing fees from traditional services, these players attempt to increase their proceeds by leasing their data to other financial market players, such as asset managers or private banks. The data thus remains the “property” of the aggregators, which allows them to restrict the usage of data for their customers.

In most sectors, companies still buy and sell hardly any data in the secondary market. Data purchases in the secondary market are of lesser importance and rather comprise generally available data, such as weather forecasts or building data. There even are companies which state a ban on selling data in their data strategy. This is due to uncertainties about data protection as well as concerns with regard to the disclosure of business secrets.

The company survey also shows that currently only a small number of Swiss companies participate in markets for trading non-personal data. However, they are expected to gain in importance. If a company offers own non-personal data in the secondary market, this data is demanded in over 40 percent of the cases by other companies, in around a quarter of the cases by authorities and in just slightly under 20 percent of the cases by research institutions, according to the survey.

2.3 Data sharing

Data sharing agreements are used when companies share self-collected data with each other. Often, they also include the management of data purchased on the secondary market. Companies have mutual targets. To achieve these targets, they grant each other access to data. Typically, they provide data to third parties via application programming interfaces (API). Supplementary services are offered to customers and the value of core services is increased (Schweitzer & Peitz, 2017, pp. 22f.).

Data sharing is particularly common for third-party devices in the mechanical engineering sector. Users encounter fewer breakdowns or shorter downtimes and for the manufacturers of the machines, the access to data improves their product development. The joint use of non-personal data

in production adds value, particularly with regard to system optimisation, tracking products and processes along the value chain, exchanging digital product properties or enabling a proof of origin (WEF, 2020, pp. 7ff.).

The expert talks in the Swiss context show that data sharing is already highly relevant in the smart farming sector. The joint use of data through open and standardised APIs requires a complex implementation process. However, machine manufacturers have apparently managed to come to an agreement with the international standard committees so that the data exchange formats in the agricultural sector are already largely standardised around the world. It is therefore possible to transfer data from one machine to another. When purchasing a machine, it is often contractually agreed which data is transferred to whom, such as for predictive maintenance, for example. But also, the farmer decides which data is transferred to whom through the data exchange portals and which apps (e.g., for animal farming) communicate with one another and how they do this.

Another data ecosystem is developing in the mobility sector, particularly with regard to the charging of electric vehicles. This process includes more than just the exchange of information between the charging status of a vehicle with the charging points: electric vehicles much rather become part of the smart power consumption process. Data related to building automation and electrification is also exchanged. Experts forecast that data sharing will gain in importance in many sectors.

2.4 Markets for data services

Ultimately, data forms the basis for the preparation of analyses and evaluations, as well as so-called data derivatives (Schweitzer & Peitz, 2017, pp. 23ff.). Demand and supply of such services determine the size of these markets for data services. The companies represented at the expert talks emphasise the importance of data derivatives compared with the mere sale of data. This is why they are primarily interested in the preparation and sale of services and insights based on their data.

On the one hand, such insights can help companies to answer their internal questions, from gaining a better understanding of the market to positioning themselves better in the market. On the other hand, they can also be used for producing services that are offered in the market. In the financial market, for example, ratings and indices are offered. As a rule, financial service providers get rated by one of the three global rating agencies, Moody's, Fitch or Standard & Poor's, thus improving their reputation. To meet the increasing demand for passive, i.e., index-based

investments, financial institutions have licence agreements with index producers such as MSCI World, FTSE 100, S&P 500, SMI or DAX. Further examples of data derivatives include analyses on the purchase behaviour as well as weather, harvest and pest forecasts in the agricultural sector.

The quality of the insights depends on the quality of the data material and specific competences (e.g., data science know-how). According to the appraisals during the expert talks, these competences are still scarce on the Swiss labour market (and elsewhere). Supplementary investments (e.g., in software or other assets) are also key. Companies and authorities are generally not so much interested in data itself, but rather in findings based on analysed data.

3. Market failure in non-personal data markets

Some experts suggest mandatory licensing to ensure that data is also available to third parties in a standard format for commercial purposes. At the same time, such mandatory licensing restricts the economic freedom of the company forced into issuing a license. Such harsh intervention must be based on the proof of market failure in markets of non-personal data. For this, we have to check for factors like asymmetric information, externalities, data-as-a-public-good, or technological peculiarities.

3.1 Asymmetric information

In information economics, we examine three broad types of asymmetric information within the scope of bi- or multilateral agreements: moral hazard, adverse selection and signalling. One contracting party has more information about the object of agreement than the other party. In the case of moral hazard, this manifests after the conclusion of the agreement (e.g., careless conduct after concluding an insurance policy). In the case of adverse selection, it manifests prior to the conclusion of the agreement (e.g., unknown condition of a second-hand car). Signalling deals with the disclosure of private information prior to contractual negotiations (see, e.g., Macho-Stadler & Pérez-Castrillo, 2001).

In the context of the data economy, the main issue is adverse selection. The persons requesting the data could be unaware of the quality of the data prior to purchase. Before using and analysing the data for the planned processes, users do not know if it, in fact, meets the required quality

standards. However, the issue of adverse selection could possibly be largely avoided by taking random samples of the data. The quality of the data can be tested by the user if he or she is able to analyse and test some of the data prior to purchasing (Schweitzer & Peitz, 2017, p. 36).

When talking about data economy, information asymmetry is often mentioned in connection with price discrimination (see, e.g., Wambach et al., 2019, p. 15; Stigler Committee, 2019, p. 60). Companies that hold comprehensive customer data (such as transport operators) can estimate their behaviour, preferences and willingness to pay and use this data for drawing up agreements as well as pricing. However, price differentiation can be both socially damaging and beneficial so that the welfare effect depends on the particularities of the situation. Companies can also use this comprehensive data for offering their customers products that better meet their preferences.

Overall, asymmetric information can probably be excluded as a possible reason for market failure in non-personal data markets.

3.2 Externalities

Externalities are another possible reason for market failure (see, e.g., Varian, 2014, ch. 35). An externality is created when the actions of one player have a direct impact on the use or production of one or several other players. It is not apparent to what extent such effects should appear when using and/or processing the data. Although the literature refers to the possibility of the existence of externalities in connection with non-personal data, hardly any specific facts are provided in this respect (Kerber, 2016, p. 10; OECD, 2015, p. 186).

Especially as the non-personal data markets are dynamic markets and a “moving target”, so to speak, and the business models are continuously changing, externalities are difficult to spot. If they do occur, the main solution is negotiations between the affected parties. Particularly in the case of positive externalities, it is attractive to internalise them in the form of agreements or integration.

3.3 Public goods

Public goods have two main properties: There is no rivalry in consumption and nobody can be excluded from consuming the public good (Jones & Tonetti, 2020).

Non-rivalry means that one person's consumption does not have a negative effect on the availability of the good for another person. Like a radio show, where listeners do not affect each other's consumption, non-personal data are obviously non-competitive in terms of consumption. An unlimited number of consumers can use the same information without this impairing the use of the data (see, e.g., Kerber, 2016). Data can be distributed via the internet to an unlimited extent at negligible cost. Therefore, non-personal data are non-rival in terms of consumption.

The second characteristic to qualify the good as a public good is non-excludability from consumption. In the case of non-personal data, consumption can often and rather easily be excluded. The actual data owner can protect the data against third-party access by simply not providing the data or by implementing corresponding encryption measures.

Non-personal data is therefore not a public good and there is also no danger of investment costs being "lost". The company that invested in the production of the data is in the position to reserve the sole right of disposal in this data and thus is able to amortise its investment costs through the sale of products and services for which the data is required. The main question when considering non-personal data is therefore not a shortfall in supply, but rather if the existing data is being used to an optimal extent.

3.4 Technological properties

As for all other goods and services, the question in the non-personal data markets is also if market power prevents efficient allocation, in other words if competition does not sufficiently discipline the players. Market power can be favoured through technological properties such as economies of scale, network effects, feedback loops.

Increasing returns to scale, for example, in the relevant area of demand result (in the long run) in only one large company or just a few large companies being able to operate in the market. This can create market power, the exploitation of which leads to inefficient allocation. The outcome will be higher prices, an insufficient product range or products that are of poorer quality than in a competitive market.

Especially in digital markets, feedback loops and network effects favour the emergence of companies that dominate the market. Network effects lead to an increasing value for the users of the platform for each additional user of the same platform. Buyers can also profit from other buyers on the platform as this attracts more buyers and vice versa. In any case, such network effects result in a platform increasing in value for users the bigger it gets.

Combined with additional factors, such as the extent of increasing scale effects, lock-in effects and system change costs, data-based companies have the tendency to become monopolies if they are successful. As the typical examples of Facebook, Alphabet (Google), Amazon, Microsoft and Apple (“FAAMA”) suggest, they can develop a winner-takes-it-all effect (see also Haucap, 2019 or Furman et al., 2019). The European Commission has found that the majority of the data in existence around the world is held by this small number of large technology companies and assumes that this reduces the incentives for the creation of data-driven companies and their growth and that it could impede innovation in the EU (European Commission, 2020, p. 3).

Overall, market failure in non-personal data markets is not likely to be expected in terms of asymmetric information, externalities or public goods. The technological properties in digital markets, on the other hand, can lead to market power constellations and therefore market entry barriers. If companies are no longer disciplined by current or potential competition, market failure occurs as a result of the exploitation of market power.

4. Market dynamics and access regulation

Put simply, data is primarily used for producing products and services that are offered in the market. The companies finance expenses such as the costs for producing, storing, processing and analysing data with revenue from the sale of these products and services. The stronger the competition for these products that are based on the data, the lower the companies’ revenues.

The more companies have access to the same data, the stronger the competition for the products and services produced with this data. If revenues no longer cover the costs for producing, storing and processing the data, it is no longer viable for companies to produce this data. Although data is non-rival in consumption, it is therefore in each company’s own interest

to use the data exclusively (private good) or to make it available only to a small group of users for exclusive use (club good).

In terms of investment incentives and data access conditions, the situation for public and private companies is therefore fundamentally different.

4.1 Special case of publicly funded data

In a blogpost for the Open Knowledge Foundation, Lämmerhirt et al. (2019) subsume data on topography, earth observation and the environment, meteorology, statistics and public traffic among the most valuable publicly financed data. Authorities, government departments and universities produce data derivatives (goods and services) from such data. However, they do not have to sell them in the market to amortise their investments in research and development, since their costs are already financed through tax revenue. It is therefore unnecessary to generate income through data exclusivity that exceeds the direct costs created from the exchange of data.

As publicly financed institutions do not need to cover their investment in data generation, processing, storage etc. by the sale of products or services, the problem of investment incentives does not arise and a static view is appropriate in this special case. As a consequence, access to data should be as broad as possible and only restricted in connection with business secrets, the protection of public safety and critical infrastructures as well as the protection of personal data (European Parliament and the Council, 2019).

In view of the investment incentives, which are fundamentally different for public institutions compared with private companies, Switzerland launched the project “Open Government Data 2014–2018” (Bundesrat, 2014) back in 2014. With the aim to promote a so-called “open data culture”, this was meant to create the basis for providing data produced by authorities and its derivatives to the general public in a suitable format. As a result of this project, the portal opendata.swiss emerged. It is a portal where the public can obtain freely usable data produced by authorities. The “Open Government Data Strategy 2019–2023” (Bundesrat, 2019) aims to further expand this portal.

4.2 Exclusive data usage and data sharing

In contrast to the case of publicly financed data, there is a regulatory trade-off for privately supplied data: should as many players as possible gain access to data, or should data-exclusivity increase investment incentives? If companies have to share their data, this decreases investment and innovation incentives with regard to the production of data or quality improvement. In return, a high volume of data by several data producers and several topics could be jointly evaluated and the potential for innovative products and services would grow strongly in the short term. In the long term, the evaluation changes. This situation shows that it is essential to take a dynamic look at market activities.

To see this, consider Clavorà Braulin and Valletti (2016), who show that a data broker has an advantage in the market if he/she uses data exclusively. If the broker had to provide the data to other companies as well at his own production costs, there would be stronger competition and therefore narrow the broker's profits from the sale of the data. From a social perspective, it might be beneficial to prohibit such exclusive relationships and therefore enable other companies to access data. However, this conclusion would only apply under the assumption that the data is exogenously given.

In a dynamic context, where the production and the supply of data is endogenously determined, the question of whether to permit exclusivity or not to have any data in the first place is therefore essential. Data must first be generated and prepared, which requires investments and labor input before it can be used. If the producer of the data can anticipate that third parties will also be allowed to use it and will compete against him in the market for the products and services that use the data as input, he may not be able to amortise the costs associated with the data production. In addition, he could also find it more attractive to hope that someone else would provide the data and that he could then free-ride on using those. The production of the data is not viable in this case. This means that companies have too little incentive to invest in data production and innovation (CMA, 2020a).

The producer of data not always knows about all the potential application ideas for it. Often individual data sources bear a considerable option value, which is not obvious or known at the point of gathering the data. In other words, the value of data can increase as soon as it starts getting used for answering new, previously unconsidered, questions (Coyle et al., 2020, p. 6). For the implementation of (potential) application ideas, it can therefore be highly useful to compile and analyse several complementary

data sets by several data producers with the aim to produce new products and services.

In a dynamic environment companies have incentives to enter into cooperations with third parties (Drexler, 2017; Open Data Institute, 2018). This was also confirmed in our expert talks. Various experts could even imagine entering into agreements with competitors for future products. However, these statements referred particularly to complementary products, or at least products that are not substitutes of the companies own products, i.e., products that do not jeopardise the amortisation of the investments in data production and processing.

By reserving exclusive rights to the data or permitting data sharing only under specific conditions, companies not only protect their investments, but can also potentially use the data for market foreclosure and possibly abuse their resulting market domination. This may make it necessary for the regulators to act by granting access rights. This raises the question under which conditions regulated data access becomes necessary from an economic perspective.

4.3 Essential Data as an economic criterion for access regulation

In terms of competition economics, the reference point for a stable market entry barrier is derived from a specific cost constellation. It requires that a production input, such as a production facility, is a “monopolistic bottleneck”, a concept from the disaggregated regulation theory established by Günter Knieps in a number of seminal articles. Accordingly, this situation requires a combination of economies of scale and irreversible costs (sunk costs) (see, e.g., Knieps, 2005, p. 103). The well-known essential facility doctrine stated in US antitrust law can be understood in this context (see in this context Knieps, 2011). It states that a facility is regarded as essential when market entry is impossible without access to this facility and a provider is unable to duplicate the corresponding facility at reasonable costs.

An incumbent firm can therefore maintain a barrier to market entry when it produces with economies of scale and if investments in the corresponding production plant are sunk costs. Sunk costs are relevant for the decisions of potential competitors as they consider taking the risk to invest capital and labor, which they may have to write-off completely should the project fail. In contrast, the company already active in the market has comparatively low costs that influence its decision-making process. Even in the case of the established provider’s production being inefficient

at the “essential facility” this does not automatically make it possible to successfully enter the market (see, e.g., Knieps, 1997, pp. 327f.).

In the EU, the refusal to supply goods and services is apparently the most frequent abuse of market power. This also includes the denial of access to an “essential facility”. European judges have therefore interpreted the misuse of obstruction for a broader scope of application than judges in the USA (Bruc, 2019, p. 183).

The European Court of Justice (ECJ) correspondingly applied the legal institution of the essential facility doctrine for the first time in 1995 in the “Magill case” with regard to non-personal data (ECJ, 1995). As Ireland had no comprehensive weekly TV programme guide back then, Magill TV Guide Ltd. planned to issue such a guide. The company was able to convince the court that the TV broadcasters Radio Telefis Eireann (RTE) and Independent Television Publications (ITP), which had based their statements on copyright protection, had to provide third parties with their weekly programme previews upon request without discrimination. If the broadcasters issued licenses for the reproduction of the previews, the license fees had to be reasonable.

The “IMS Health case” (ECJ, 2004) is another important decision of the ECJ on the topic of data access based on the logic of an essential facility. The parties considered the production and sale of market reports for the sale of products such as medicinal drugs. The claimant produced reports for pharmaceutical companies based on building block structures protected by intellectual property rights. The respondent (NDC Health), who also planned to provide market reports in the medicinal sector, was ultimately able to obtain a license to this building block structure from IMS Health through the court.

Interoperability can also be qualified in regulatory terms as data access, and a deficiency of the same as an abuse of market power (Andreangeli, 2009; Hoffmann & Otero, 2020; OECD, 2018). A key-case in this respect is the obligation placed on Microsoft by the European Commission to disclose the interface specifications of various software products so that operating systems of competing companies can communicate unrestrictedly with client PCs and workgroup servers that use Microsoft Windows.

The European Commission sees the need for regulatory intervention with regard to non-personal data particularly in the presence of signs for a market concentration. The European Commission claims that under “special circumstances”, the granting of data access under fair, reasonable, appropriate and non-discriminatory conditions (FRAND-conditions) should be made mandatory, but does not suggest a specific form of access regulation (see European Commission, 2020, p. 16).

The aforementioned special circumstances translate to the test if the data is essential data, similarly to physical essential facilities. The conditions for essential data can be summarised as follows (Abrahamson, 2014, pp. 104f.; Duch-Brown et al., 2017, p. 21):

- The dominant company in the market is able to grant data access but refuses to do so to third parties.
- Competition for the services based on this data is impossible without access to this data.
- Third parties are unable to duplicate the data and/or create a comparable database.

Once these conditions have been tested, it can be triaged if the data access refusal constitutes misconduct of a dominant company in the market or a legitimate economic concern of the affected company. The right to access non-personal data should only be granted if the data satisfy the properties of essential data. In this case, the conditions for accessing the data must be balanced so that incentives for quality improvements and technological progress are maintained, particularly with regard to the continued production of this non-personal data.

5. Implications from the perspective of competition and regulatory economics

Data markets are developing rapidly and the corresponding business models are changing continuously. Currently, companies are not primarily interested in purchasing data per se, but in the findings from the data. However, access to essential data might be impeded or denied by a dominant firm. If such business practices do not occur systematically but ad hoc, they can be effectively disciplined with the help of competition laws.

Only a minority of the surveyed companies which request data from third parties are currently mentioning difficulties, such as excessive license fees, disadvantageous contractual terms and conditions, or technical barriers. As the market is not generally failing, it is not indicated to effect access to essential data through legally binding, mandatory licenses under FRAND-conditions. This applies regardless of such licenses being based on intellectual property rights or competition law (IGE, 2021, pp. 24f.). Even for the European Commission, such licenses look like they no longer are a priority for the granting of access to non-personal data, especially as they are no longer mentioned in the European data strategy (European Commission, 2020).

Competition authorities primarily operate through ex post measures, for example, by imposing fines after an infringement has been established. They do normally act ex ante as it would be the case with sector-specific regulatory measures. However, in order to deal with individual competition policy issues, Switzerland also has at its disposal tried-and-tested instruments such as announcements that can have an ex ante effect (see Competition Commission, 2015a, Art. 14 and 17; Competition Commission, 2015b). It generally could also use such instruments for questions regarding access to non-personal data. However, the introduction of the market survey instrument could also be investigated. Such instruments are already known to the UK (CMA, 2020a) or the Netherlands (Authority for Consumers and Markets, 2016) and are discussed by the European Commission (CMA, 2020b; see also Höhn, 2020). Should the Competition Commission find negative effects on competition during such market survey, it would be able to implement its own measures or propose measures for strengthening competition to other authorities (regulators) or the legislator.

Based on the practice of the European Commission, the right to data access, on the other hand, should only be granted by sector-specific regulatory measures, if a market failure is identified in the corresponding sector, if it cannot be remedied with competition law alone, and if the legitimate interests of the data owner are accounted for. In this context, selected sector-specific legal provisions on data access are already being issued in the EU to remedy any market failure that has been discovered (European Commission, 2020, p. 4).

6. Conclusion

Producing, collecting and storing data is cost-intensive, so regulated access to non-personal data from private companies is likely to generally thwart the creation of good data sets. Unlike in the case of publicly financed data, companies have to finance their investments from the sale of products and services, which often requires the exclusive use of data. The investment incentives erode if this exclusive access is no longer available. From a dynamic perspective, it is therefore important that private companies can decide on data access and/or the exclusion of third parties on a contractual basis.

Under certain conditions, companies have incentives to share data with other companies. However, if substitutes for a company's own products and services can be produced through the access to data, companies are going to refuse to grant data access. Nevertheless, the refusal of data access

does not necessarily constitute inefficient conduct or even misconduct. Regulatory data access for third parties is only justified for so-called “essential data”. Access regulation (such as mandatory licensing) therefore requires for proof of data to be essential in each specific case.

If market power remains stable, sector-specific regulations can be used in order to prevent the abuse of market power (e.g., in the case of access to measuring data in the electricity sector). In all other cases, access regulation should be implemented through competition law. As competition law is primarily applied *ex post* and using rather tough measures (fines and “remedies”), it also has to be investigated to what extent competition authorities could play a more active and constructive role through (new) *ex ante* instruments.

With regard to potential policy recommendations, the introduction of instruments such as the “market investigations” of the competition authorities in the UK and Netherlands may be worth investigating. The key feature of market surveys is that the burden of proof is less than during the investigation of a dominant market position. From the point of view of competition policy, this leaves for future research the question, under which circumstances such market surveys would be justified, particularly in view of the predictability of legal decisions.

References

- Abrahamson, Z. (2014). Essential Data. *The Yale Law Journal*, 124(3), 867–881. <https://www.jstor.org/stable/43617042>.
- Andreangeli, A. (2009). Interoperability as an “Essential Facility” in the Microsoft Case Encouraging Competition or Stifling Innovation. *European Law Review*, 4, 584–611.
- Authority for Consumers and Markets (2016). *Prioritization of enforcement investigations by the Netherlands Authority for Consumers and Markets* (Dutch Government Gazette No 14564). Authority for Consumers and Markets.
- Bruc, É. (2019). Data as an Essential Facility in European Law: How to Define the „Target“ Market and Divert the Data Pipeline? *European Competition Journal*, 15(2–3), 177–224. <https://doi.org/10.1080/17441056.2019.1644576>.
- Bundesrat (2014). Open-Government-Data-Strategie Schweiz 2014–2018. *Bundesblatt*, 2014, 3493–3506. <https://www.fedlex.admin.ch/eli/fga/2014/747/de>.
- Bundesrat (2019). Strategie für offene Verwaltungsdaten in der Schweiz 2019–2023 (Open-Government-Data-Strategie, OGD-Strategie). *Bundesblatt*, 2019, 879–894. <https://www.fedlex.admin.ch/eli/fga/2019/125/de>.

- Clavorà Braulin, F. & Valletti, T. (2016). Selling Customer Information to Competing Firms. *Economics Letters*, 149, 10–14. <https://doi.org/10.1016/j.econlet.2016.10.005>.
- CMA (2020a). *Online Platforms and Digital Advertising Market Study* (Final Report). Competition and Markets Authority. https://assets.publishing.service.gov.uk/media/5fa557668fa8f5788db46efc/Final_report_Digital_ALT_TEXT.pdf.
- CMA (2020b). *The CMA's response to the European Commission's consultations in relation to the Digital Services Act package and New Competition Tool*. Competition and Markets Authority. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/917455/CMA-response_to_DSA_and_NCT_consultations.pdf.
- Competition Commission (2015a). Wettbewerbskommission. *Bekanntmachung über die wettbewerbsrechtliche Behandlung von vertikalen Abreden im Kraftfahrzeugssektor vom 29. Juni 2015* (Stand am 9. September 2019).
- Competition Commission (2015b). Wettbewerbskommission. *Erläuterungen der Wettbewerbskommission zur Bekanntmachung über die wettbewerbsrechtliche Behandlung von vertikalen Abreden im Kraftfahrzeugssektor vom 29. Juni 2015* (Stand am 9. September 2019).
- Coyle, D., Diepeveen, S., Wdowin, J., Kay, L. & Tennison J. (2020). *The Value of Data – Policy Implications* (Main Report). Bennett Institute for Public Policy, Open Data Institute & Nuffield Foundation. https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value_of_data_Policy_Implications_Report_26_Feb_ok4noWn.pdf.
- Drexl, J. (2017). Designing Competitive Markets for Industrial Data – Between Propertisation and Access. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 8, 257–292.
- Duch-Brown, N., Martens, B. & Mueller-Langer, F. (2017). *The Economics of Ownership, Access and Trade in Digital Data* (JRC Digital Economy Working Paper 2017–01). European Commission. <https://dx.doi.org/10.2139/ssrn.2914144>.
- ECJ (1995). Judgement of 6 April 1995. *Radio Telefís Éireann (RTE) and Independent Television Publications Ltd (ITP) v Commission of the European Communities*, Joint Cases C241/91 P and C242/91 P (ECLI:EU:C:1995:98).
- ECJ (2004). Judgement of 29 April 2004. *IMS Health*, C-418/01 (ECLI:EU:C:2004:257).
- European Commission (2018). *Commission Staff Working Document: Guidance on sharing private sector data in the European data economy: Accompanying the document: Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions: "Towards a common European data space"* (SWD(2018) 125 final). European Commission. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018SC0125>.

- European Commission (2019a). *Communication from the Commission to the European Parliament and the Council: Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union* (COM(2019) 250 final). European Commission. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A250%3AFIN>.
- European Commission (2019b). *SME Panel Consultation on B2B Data Sharing Principles and Guidance: Report on the Results* (Final report). European Commission. <https://digital-strategy.ec.europa.eu/en/library/sme-panel-consultation-b2b-data-sharing>.
- European Commission (2020). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European strategy for data* (COM(2020) 66 final). European Commission. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>.
- European Parliament and the Council (2018a). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L 119, 1–88. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.
- European Parliament and the Council (2018b). Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union. *Official Journal of the European Union*, L 303, 59–68. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32018R1807>.
- European Parliament and the Council (2019). Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast). *Official Journal of the European Union*, L 172, 56–83. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L1024&qid=1636393439048>.
- Frontier Economics (2021). *Increasing Access to Data across the Economy. A report prepared for the Department for Digital, Culture, Media and Sport* (Policy paper). Department for Digital, Culture, Media and Sport. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974532/Frontier-access_to_data_report-26-03-2021.pdf.
- Furman, J., Coyle, D., Fletcher, A., McAuley, D. & Marsden, P. (2019). *Unlocking Digital Competition. Report of the Digital Competition Expert Panel* (Independent report). HM Treasury. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf.
- Haucap, J. (2019). Competition and Competition Policy in a Data Driven Economy. *Intereconomics*, 54, 201–208. <https://doi.org/10.1007/s10272-019-0825-0>.
- Hoffmann, J. & Otero, B. G. (2020). *Demystifying the role of data interoperability in the access and sharing debate* (Research paper No. 20–16). Max Planck Institute for Innovation & Competition. <http://dx.doi.org/10.2139/ssrn.3705217>.

- Höhn, T. (2020). Challenges in Designing and Implementing Remedies in Innovation Intensive Industries and the Digital Economy. In D. Gerard & A. Komninou (Eds.), *Remedies in EU Competition Law: Substance, Process and Policy* (pp. 121–148). Alphen aan den Rijn, NL: Wolters Kluwer.
- IGE (2021). *Zugang zu Sachdaten in der Privatwirtschaft*. Institut für Geistiges Eigentum. https://www.ige.ch/fileadmin/user_upload/recht/gesellschaft/d/20210301_Bericht_IPI_Zugang_zu_Sachdaten_in_der_Privatwirtschaft.pdf.
- Jones, C. I. & Tonetti, C. (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110(9), 2819–2858. <http://dx.doi.org/10.1257/aer.20191330>.
- Kerber, W. (2016). A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis. *Gewerblicher Rechtsschutz und Urheberrecht. Internationaler Teil*, 11, 989–998. <https://ssrn.com/abstract=2858171>.
- Knieps, G. (1997). Phasing out Sector-Specific Regulation in Competitive Telecommunications, *Kyklos*, 50(3), 325–339. <https://doi.org/10.1111/1467-6435.00019>.
- Knieps, G. (2005). *Wettbewerbsökonomie: Regulierungstheorie, Industrieökonomie, Wettbewerbspolitik* (2., überarbeitete Auflage). Berlin & Heidelberg, DE: Springer.
- Knieps, G. (2011). The Three Criteria Test, the Essential Facilities Doctrine and the Theory of Monopolistic Bottlenecks. *Intereconomics*, 46(1), 17–21.
- Macho-Stadler, I. & Pérez-Castrillo, J.D. (2001). *An Introduction to the Economics of Information: Incentives and Contracts*. Oxford, UK: Oxford University Press.
- OECD (2015). *Data-Driven Innovation: Big Data for Growth and Well Being*. OECD Publishing. <http://dx.doi.org/10.1787/9789264229358-en>.
- OECD (2018). *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*. OECD Publishing. <https://doi.org/10.1787/276aac-a8-en>.
- Open Data Institute (2018). *What organisations need in order to share more data*. Open Data Institute. <https://theodi.org/article/what-organisations-need-in-order-to-share-more-data-our-research/>.
- Lämmerhirt, D., Chrzanowski, P. & van der Waal, S. (2019, January 16). What data counts in Europe? Towards a public debate on Europe's high value data and the PSI Directive. *Open Knowledge Foundation*. <https://blog.okfn.org/2019/01/16/what-data-counts-in-europe-towards-a-public-debate-on-europes-high-value-data-and-the-psi-directive/>.
- Schneider, Y., Zenhäusern P. & Vaterlaus, S. (2020). *Märkte für Sachdaten in der Schweiz: Ökonomische Analyse der Funktionsweise von Märkten für Sachdaten in der Schweiz unter Einbezug der Unternehmenslandschaft. Studie im Auftrag des Instituts für Geistiges Eigentum (IGE)*. Polynomics. https://www.ige.ch/fileadmin/user_upload/recht/gesellschaft/d/201113_Polynomics_Schlussbericht-Sachdaten.pdf.
- Schweitzer, H., & Peitz, M. (2017). *Datenmärkte in der digitalisierten Wirtschaft: Funktionsdefizite und Regelungsbedarf?* (Discussion Paper No. 17–043). ZEW – Zentrum für Europäische Wirtschaftsforschung. <http://ftp.zew.de/pub/zew-docs/dp/dp17043.pdf>.

- Stigler Committee (2019). *Stigler Committee on Digital Platforms: Final Report*. Chicago Booth Stigler Center for the Study of the Economy and the State. <https://www.chicagobooth.edu/-/media/research/stigler/pdfs/digital-platforms---committee-report---stigler-center.pdf>.
- UVEK (2019). *Bericht zu den Empfehlungen der Expertengruppe zur Zukunft der Datenbearbeitung und Datensicherheit: Kenntnisnahme und weiteres Vorgehen*. Schweizerische Eidgenossenschaft – Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK. <https://www.news.admin.ch/news/message/attachments/58796.pdf>.
- Varian, H. R. (2014). *Intermediate Microeconomics with Calculus* (1st Edition). New York, NY, US: WW Norton & Co.
- Wambach, A., Schweitzer, H. & Schallbruch, M. (2019). *Ein neuer Wettbewerbsrahmen für die Digitalwirtschaft: Bericht der Kommission Wettbewerbsrecht 4.0*. Bundesministerium für Wirtschaft und Energie. https://www.bmwi.de/Redaktion/DE/Publikationen/Wirtschaft/bericht-der-kommission-wettbewerbsrecht-4-0.pdf?__blob=publicationFile&v=14.
- Weber, R. H. & Thouvenin, F. (2018). Dateneigentum und Datenzugangsrechte – Bausteine der Informationsgesellschaft? *Zeitschrift für Schweizerisches Recht*, 137(1), 43–74.
- WEF (2020). *Share to Gain: Unlocking Data Value in Manufacturing*. White Paper. World Economic Forum in collaboration with Boston Consulting Group. https://www3.weforum.org/docs/WEF_Share_to_Gain_Report.pdf.