

Zeitschrift für

# Vergleichende Rechts- wissenschaft

Archiv für Internationales Wirtschaftsrecht

## 3

### **Einfluss außerrechtlicher Faktoren auf Recht und Rechtsvergleichung**

**Christian Kirchner und Uwe Kischel**

Einfluss außerrechtlicher Faktoren auf Recht und  
Rechtsvergleichung (269)

**Kilian Bälz**

Islamisches Kreditwesen – Religion, Wirtschaft und Recht  
im Islam (272)

**Walter Homolka**

Der Einfluss nationalen Rechts auf jüdisches Denken am  
Beispiel des Jüdischen Eherechts (293)

**Prakash A. Shah**

Comparatively Indian: Living with legal plurality (314)

**Oliver Lepsius**

Der Einfluss des ökonomischen Denkens auf die  
Rechtsvergleichung (327)

**Katharina Pistor**

Statistische Rechtsvergleichung: Eine kritische  
Bestandsaufnahme (348)

**Axel Tschentscher**

Rechtsvergleichung und empirische Forschung (362)

# Rechtsvergleichung und empirische Forschung

Axel Tschentscher\*

ZVglRWiss 109 (2010) 362–377

## I. Überblick

Eine Empirie, die sich nicht der sozialwissenschaftlichen Methodik bedient, schwingt in der Rechtsvergleichung immer schon mit, denn neben den Rechtsnormen ist die Rechtspraxis ihr Gegenstand. Meistens führt diese methodisch unflektierte Wirklichkeitsbetrachtung zu intuitiv plausiblen Annahmen. Einzelfällen aus der Rechtsprechung oder Erfahrungsberichten von Länderexperten veranschaulichen den Normbefund durch den Kontext der Rechtskultur. Die Frage ist also nicht, ob *überhaupt* Empirie in der Rechtsvergleichung verwendet wird, sondern vielmehr, ob die methodisch stringendere Empirie der sozialwissenschaftlichen Forschung für die normative Rechtsvergleichung einen Zusatznutzen bringt.

Was zeichnet die Empirie unter den außerrechtlichen Faktoren, denen diese Tagung gewidmet ist, besonders aus? Man kann die externen Einflüsse auf die normative Rechtsvergleichung unterteilen in 1. nichtrechtlich-normative, 2. analytische und 3. empirische.

### 1. Nichtrechtlich-normative Faktoren

Unter den nichtrechtlich-normativen Faktoren, die in der Praxis relevant werden, sind vor allem diejenigen der Religion zu benennen. Das staatliche Recht kann sich religiösen Normen öffnen, indem es der Religion innerhalb des rechtlichen Rahmens einen Zuständigkeitsbereich einräumt. Das ist beispielsweise der Fall, wenn für die Anerkennung der Schließung oder Auflösung einer Ehe auf innerreligiöse Regelungen verwiesen wird. Man kann insoweit von einer rechtlichen Pluralität im schwachen Sinne sprechen. Als Ergebnis kommt es zu einer unechten, weil kompetenziell aufgelösten Konkurrenz zwischen Normensystemen.

Eine rechtliche Pluralität im starken Sinne findet sich demgegenüber dort, wo staatliches Normensystem und religiöses Normensystem konkurrierende Geltungsansprüche erheben, ohne aufeinander abgestimmt zu sein – weder re-

---

\* Prof. Dr. Axel Tschentscher, LL.M. (Cornell), attorney-at-law (New York), Lehrstuhl für Staatsrecht, Rechtsphilosophie und Verfassungsgeschichte an der Universität Bern, Bern Center for Interdisciplinary Deliberation Studies.

kurriert das staatliche Recht auf innerreligiöse Bestimmungen noch anerkennt das religiöse Recht den Geltungsvorrang der jeweiligen staatlichen Ordnung. Die Pluralität der im Ergebnis wirkenden Normen ergibt sich dann daraus, dass jede der Ordnungen nur unvollständig verwirklicht wird, beispielsweise wenn sich die Scheidungsfolgenregeln eigentlich aus dem staatlichen Recht ergeben, faktisch aber in einzelnen Religionsgruppen deren religiöse Gebote zur Anwendung kommen. Im Ergebnis führt das zu einer echten Konkurrenz zwischen Normensystemen.

## 2. Analytische Faktoren

Bei der analytischen Forschung geht es nicht um Beobachtungen der Wirklichkeit, sondern um Berechnungen. In Anlehnung an *Kant* könnte man sagen, dass es sich um Erkenntnisse handelt, die *vor jeder Erfahrung* liegen. Die analytische Zentraldisziplin ist die Mathematik. Das mathematische Teilgebiet, das sich mit der Analyse sozialen Verhaltens beschäftigt, ist die Spieltheorie. Mit ihr wird versucht, ein rationales Entscheidungsverhalten (*rational choice*) in Situationen mit mehreren Akteuren zu berechnen. Meist handelt es sich um Konfliktsituationen. Das wohl bekannteste Beispiel für eine solche Situation ist das Gefangenendilemma.

Der Berner Wirtschaftswissenschaftler *Winand Emons* hat gemeinsam mit dem in Berkeley arbeitenden Rechtssoziologen *Robert Cooter* beispielsweise berechnet, wie man die Wahrheitsfindung in Prozessen verbessern könnte.<sup>1</sup> Ausgangs- und Vergleichspunkt ihres Kalküls war ein Zustand, in dem Zeugen und Sachverständige für unpräzise Aussagen nur sehr selten mit Konsequenzen rechnen müssen. Das ist in den meisten Rechtsordnungen heute der Normalfall. Mit ihrer Berechnung versuchen *Emons* und *Cooter* zu beweisen, dass man die prozessuale Wahrheitsfindung insgesamt begünstigen kann, wenn man eine verschuldensunabhängige Sanktion bei objektiv falschen Aussagen vorsieht. Angesichts verschuldensunabhängiger Konsequenzen wären Zeugen und Sachverständige von vornherein vorsichtiger und im Ergebnis gewissenhafter mit ihren Behauptungen. Darunter leidet zwar die Aussagebereitschaft, aber insgesamt ergibt sich nach dieser Analyse ein Vorteil für die Wahrheitsfindung.

Wie kann analytische Forschung dieser Art für die Rechtsvergleichung bedeutsam werden? Für die Vergleichung bietet sie ein *tertium comparationis*, an dem die realen Rechtsordnungen gemessen werden. Wenn man dem in der Rechtsvergleichung verbreiteten Ideal des neutralen Standpunkts anhängt,<sup>2</sup>

1 *Winand Emons/Robert D. Cooter*, Truth-Revealing Mechanisms for Courts, in: SSRN Working Paper Series Nr. 233090, May 2000.

2 Das ist bisher die überwiegende Grundeinstellung in der rechtsvergleichenden Methodik; für Gründe gegen das Ideal der Neutralität in der Rechtsvergleichung siehe *Axel Tschentscher*, Dialektische Rechtsvergleichung. Zur Methode der Komparistik im öffentlichen Recht, JZ 2007, 807, 812 ff.

dann hat das analytische Referenzmodell zudem den Vorteil, noch gar nicht verwirklicht worden zu sein. Das Modell ist ein reiner Vorschlag *de lege ferenda* und als solcher national unverdächtig, weil es eine abstrakte Verhaltensrationalität beschreibt, an der sich die bisherigen Normen der verschiedenen Rechtsordnungen mehr oder weniger stark orientieren. Es gibt kein Vorbildland, wir haben vielmehr einen quasi-objektiven außerrechtlichen Maßstab, um die Rechtsordnungen zu messen.

### 3. Empirische Faktoren

Bei den empirischen Faktoren ist das ganz anders. Studien über die Wirklichkeit des Rechts betreffen immer konkrete Rechtsordnungen. Sie sind dadurch von Anfang an national gefärbt und stehen unter dem Verdacht, nichts anderes als die Besonderheiten eines einzelnen Rechtssystems auszudrücken. Man kann sagen, dass Rechtsempirie *prima facie* nicht vergleichbar ist. Es ist jeweils eine methodische Hürde zu überwinden, und dafür gibt es unterschiedliche Vorgehensweisen. Im Folgenden werden drei dieser Vorgehensweisen vorgestellt.

## II. Leistungsbezogene empirische Forschung

### 1. Empirisches Benchmarking

In ganz unverdächtig, weil länderneutraler Weise kann man empirische Forschung ähnlich wie analytische Forschung im Rechtsvergleich einsetzen, wenn man sie als *Benchmarking* verwendet. Das setzt allerdings voraus, dass die Vergleichung sich auf ein Sachgebiet bezieht, in dem sich die Rechtsordnungen über eine ideale Referenz bereits einig sind. Das insoweit geeignetste Gebiet ist die durch Kriminologie gestützte Strafrechtswissenschaft. Die allgemeine Zielsetzung ist bekannt: niedrige Kriminalitätsraten, hohe Aufklärungsquoten, wenig Freiheitsstrafen, geringe Rückfallquoten und alles dies mit möglichst geringem Aufwand. Wenn man solche Erfolgsbedingungen für das Recht formulieren kann, dann wirkt empirische Forschung sogar als Dreh- und Angelpunkt der Rechtsvergleichung. Die in der einen Rechtsordnung neu eingeführten Normen und Maßnahmen werden von den anderen Rechtsordnungen unmittelbar als Modellversuch wahrgenommen und können, wenn sie empirisch betrachtet Erfolg haben, als neue Referenzordnung dienen. In der Wirtschafts- und Politikwissenschaft spricht man insoweit heute von *best practices*.

### 2. Einige Beispiele

#### a) Schwarzfahren in Zürich (CH)

Zur Referenzfunktion ein erstes Beispiel: Jüngst hat im *European Journal of Criminology* ein Zufallsergebnis aus der empirischen Studie des Kriminolo-

gen *Martin Killias* an der Universität Zürich für Aufsehen gesorgt.<sup>3</sup> Mit der Studie sollten die vom Zürcher Verkehrsverbund (ZVV) vor einigen Jahren neu eingeführten Abendkontrollen ab 21 Uhr ausgewertet werden. Diese Kontrollen wurden praktisch kostenneutral durch die Zugbegleiter durchgeführt, die zur Verbesserung des Sicherheitsgefühls der Fahrgäste ohnehin nachts in den Zügen patroullierten. Überraschend war nicht, dass sich damit die abendliche Schwarzfahrerquote innerhalb von drei Jahren von 3,5 auf 1 Prozent reduzieren ließ, sondern vielmehr, dass eine signifikante Verminderung der Quote auch in den Zeiten ohne Kontrollen eintrat.<sup>4</sup> Offenbar ist man in Zürich auf eine gleichermaßen schonende wie wirksame Maßnahme gestoßen. Obwohl sich die Sanktionen für das Schwarzfahren nach Art und Durchsetzung von Land zu Land deutlich unterscheiden, insbesondere die Staatsanwaltschaft zu ganz unterschiedlichen Zeitpunkten eingeschaltet wird, ist diese Art empirischer Forschung zu Recht in einer europaweit gelesenen englischsprachigen Fachzeitschrift publiziert worden. Wegen der überall gleichen Ziele kann die Erfahrung aus Zürich auch in anderen Städten und Ländern als Referenzmodell dienen.

#### b) Einbrüche in Oldham (GB)

Es gibt also Rechtsgebiete, in denen die Empirie geradezu selbstverständlich in eine internationale Vergleichung von Recht und Rechtsdurchsetzung mündet. Das ist allerdings nicht auf vielen Gebieten und selbst in diesen nicht bei allen empirischen Studien der Fall.

Nochmals zur Kriminologie: Die dramatische Reduktion von Hauseinbrüchen, die in Großbritannien durch das so genannte *Alley-Gating* erreicht werden konnte<sup>5</sup> – das ist eine seitliche Absperrung der Zugangspfade hinter Reihenhäusern – lässt sich schon deshalb nicht auf andere Länder übertragen, weil die britische Bauform der Rücken-an-Rücken-Reihenhäuser (*back-to-back terraced houses*) in anderen Teilen Europas eher unüblich ist.

#### c) Aufwandsabhängige Anwaltshonorare (USA)

Als besonderes Hindernis für ein rechtswissenschaftliches *Benchmarking* wirkt sich der Atlantik aus. Die Vereinigten Staaten von Amerika sind reich an Rechtsempirie und pflegen diese sogar an den Rechtsfakultäten. Ein Großteil der rechtswissenschaftlichen Empirie in den USA beschäftigt sich allerdings

3 *Martin Killias/David Scheidegger/Peter Nordenson*, The Effects of Increasing the Certainty of Punishment: A Field Experiment on Public Transportation, in: *European Journal of Criminologie* 6 (2009), S. 387–400. Zur öffentlichen Beachtung siehe die Presseberichterstattung „Abendliche Billettkontrollen wirken sich auch am Tag aus. Weniger Schwarzfahrer auf der S-Bahn“, in: *NZZ* Nr. 184 vom 12. 8. 2009, S. 43.

4 *Killias/Scheidegger/Nordenson* (Fn. 3), S. 387, 394 f., 398.

5 *John Haywood/Paula Kautt/Andrew Whittaker*, The Effects of ‘Alley-Gating’ in an English Town, *Euro.J.Crim.* 6 (2009), 361, 367 f., 376.

mit Besonderheiten des dortigen Rechtssystems, die ihren Weg – jedenfalls bisher – nicht nach Europa gefunden haben.

Das gilt für *punitive damages* ebenso für *forum shopping*, *class action* und *contingency fees*. Beispielsweise gibt es eine Studie, die der Rechtssoziologe Theodore Eisenberg an der Cornell Law School zum Anwaltshonorar bei *class action* erstellt hat und die ausgesprochen interessant ist, weil sie zeigt, wie die von den Gerichten eigentlich bezweckte Aufwandsabhängigkeit des Honorars – die *lodestar calculation* – die prozessuale Wirklichkeit gerade nicht bestmöglich wiedergibt.<sup>6</sup> Weil aber das Sammelklageverfahren der USA nicht unmittelbar mit europäischen Gruppenklagen – beispielsweise mit der deutschen Streitgenossenschaft, der Verbandsklage, dem neuen Kapitalanlegermusterverfahren oder auch der britischen *group litigation*<sup>7</sup> – vergleichbar ist, und weil erfolgsabhängige Honorare nach Höhe der erzielten Schadensersatzsumme in Europa weithin unüblich sind, lässt sich aus solchen Studien rechtsvergleichend nicht viel gewinnen. Selbst in Schweden, wo Sammelklage und Erfolgshonorar kombiniert werden können, sind die in Amerika möglichen prozentualen Beteiligungen der Anwaltschaft bis zu 30 % nicht möglich.<sup>8</sup>

#### d) Richterliche Amtsdauer?

Abgesehen von solchen Detaildifferenzen scheitert das empirische *Benchmarking* in den meisten Rechtsgebieten daran, dass in ihnen die gemeinsamen Kriterien für erfolgreiche Funktionserfüllung fehlen. So lässt sich für die Funktion „unparteiische Rechtsprechung“ zwar das Kriterium „unabhängige Richterschaft“ noch als konsentiert ansehen. Welche messbaren Faktoren in der Wirklichkeit allerdings diese Unabhängigkeit garantieren, ist nur schwer festzulegen, weil sie vom Kontext der jeweiligen Rechtsordnung abhängig ist. Man kann beispielsweise nicht die durchschnittliche richterliche Amtsdauer als Indikator für richterliche Unabhängigkeit verwenden. Wenn nämlich in einer Rechtsordnung eine Amtsperiode von 12 Jahren ohne Wiederwahl vorgeschrieben ist, dann würde es höchstens gegen die Unabhängigkeit sprechen, wenn die tatsächliche Amtsdauer unter diesen 12 Jahren liegt. Man darf aber nicht dieser Rechtsordnung eine geringere Unabhängigkeit attestieren als einer anderen, in der keine solche Begrenzung existiert, so dass sich die durchschnittliche Amtsdauer ohne Weiteres bei mehr als 20 Jahren einpendelt. Es könnte im Gegenteil so sein, dass die lange Amtsdauer nur durch ständige Gefälligkeit gegenüber der Exekutive erreicht wird, wohingegen die feste Amtsperiode gerade die Freiheit eröffnet, unabhängig zu agieren. Die Beurteilung der *de facto* Unabhängigkeit hängt also von den Garantien der *de lege* Unab-

6 Theodore Eisenberg/Geoffrey P. Miller, Attorneys Fees in Class Action Settlements: An Empirical Study, J.Empir.Leg.Stud. 1 (2004), 27, 63 f.

7 Zu den Einzelheiten Peter Mattil/Vanessa Desoutter, Die europäische Sammelklage. Rechtsvergleichende und EU-rechtliche Betrachtungen, WM 2008, 521, 522 ff.

8 Mattil/Desoutter (Fn. 7), S. 523 f.

hängigkeit ab. Es gibt keine schlichte Vergleichbarkeit mehr auf der Ebene der unmittelbaren Beobachtung.

### 3. Fazit

Als knappes Zwischenergebnis lässt sich festhalten, dass empirische Forschung im Sinne eines *Benchmarking* der Rechtsordnungen eingesetzt werden kann, sich der Anwendungsbereich dieser Methode allerdings in engen Grenzen hält.

## III. Systemunabhängige empirische Forschung

Damit bleibt die Frage, wie sich Empirie, die zwischen den Rechtsordnungen nicht unmittelbar vergleichbar ist, für die Rechtsvergleichung gleichwohl methodisch fruchtbar machen lässt.

### 1. Klassifizierung nach Erkenntnisinteressen

Unterscheiden sollte man zunächst zwischen *systemunabhängiger* und *systemspezifischer* empirischer Forschung. Als systemunabhängige empirische Forschung kann diejenige Untersuchung der Wirklichkeit gelten, die ihr Erkenntnisinteresse ungeachtet der Besonderheiten der beteiligten Rechtsordnungen formuliert. Letztlich geht es um Fragestellungen, die in *jeder* Rechtsordnung gleichermaßen sinnvoll sind. Als erstes Beispiel wird hier der Messung richterlicher Unabhängigkeit nachgegangen.

Demgegenüber ist *systemspezifische* empirische Forschung dadurch gekennzeichnet, dass ihr Erkenntnisinteresse auf die Instrumente einzelner Rechtsordnungen gerichtet ist. Es geht also um Fragestellungen, die *nicht in jeder* Rechtsordnung gleichermaßen sinnvoll sind. Als Beispiel kann die Messung der Diskursqualität in unterschiedlich konzipierten Staatsorganen gelten.

Die systemspezifische empirische Forschung verdient eine getrennte Behandlung, weil sie ihr Erkenntnisinteresse zwar näher an der rechtswissenschaftlichen Disziplin formuliert, dadurch aber auch zusätzliche Vermittlungsschritte nötig macht, um eine Vergleichbarkeit herzustellen. Die Differenzierung zwischen systemunabhängiger und systemspezifischer empirischer Forschung ist allerdings weder besonders trennscharf durchzuführen noch kann sie als eine zwingende oder gar einzigartige Unterscheidung gelten.

## 2. Richterliche Unabhängigkeit – *de facto* und *de iure*

### a) Fragestellung

Als Beispiel für eine systemunabhängige Fragestellung, die durch empirische Forschung zu untersuchen ist, kann hier die Messung der richterlichen Unabhängigkeit wieder aufgegriffen werden. Unter den zahlreichen Arbeiten dazu zeichnen sich diejenigen von *Lars Feld* und *Stefan Voigt* dadurch aus, dass sie die Kontextabhängigkeit der beobachtbaren Daten methodisch reflektieren. Die Notwendigkeit dazu hat sich schon am Beispiel der richterlichen Amtsdauer gezeigt. Die durchschnittliche Amtsdauer hat allein keine Aussagekraft, sondern sie muss zum normativen Rahmen der jeweiligen Rechtsordnung in Beziehung gesetzt werden. *Feld* und *Voigt* tragen dem Rechnung, indem sie zwischen *de facto*- und *de iure*-Unabhängigkeit unterscheiden und beide empirisch untersuchen. Letztlich aussagekräftig ist, was *de facto* als Unabhängigkeit resultiert, also eine systemunabhängige Frage. Aber die Antworten lassen sich nicht einfach mit einem empirischen *Benchmarking* anerkannter Kriterien finden, sondern es muss eine *normativ informierte Empirie* angewendet werden.

### b) Vorgehen

Zur grundlegenden Vereinfachung wird für jedes Land die Richterschaft des höchsten Gerichts analysiert, gleich ob es ein reines Verfassungsgericht, ein oberstes Appellationsgericht oder ein Gericht mit beiden Funktionen ist.<sup>10</sup> Die Untersuchung beruht auf einem Fragebogen, der von Experten aus insgesamt 71 Ländern ausgefüllt wurde. Als auswertbar erwiesen sich danach die Daten zu 66 Ländern.

Für *de iure*-Unabhängigkeit wird die gesetzliche Regelung nach 23 Charakteristika beurteilt, die zu zwölf Variablen kumuliert sind. Beispielsweise setzt sich die Variable „Stabilität der Unabhängigkeitsgarantie“ danach zusammen, ob die Garantie in der Verfassung steht und wie stark eine Verfassungsänderung erschwert wird. Andere Variablen betreffen beispielsweise das Verfahren der Richterberufung und die Transparenz der Tätigkeit.

Für *de facto*-Unabhängigkeit werden acht Variablen bestimmt, unter anderem auch die tatsächliche durchschnittliche Amtsdauer, allerdings mit normativen Korrekturen: Wenn die tatsächliche Amtsdauer von der normativ zu erwartenden abweicht, zum Beispiel weil Richterinnen oder Richter freiwillig vorzeitig auf ihr Amt verzichten, dann wird die Variable mit Null kodiert. Dasselbe gilt erst recht, wenn eine Richterin oder ein Richter gegen den eigenen Willen vorzeitig entlassen wurde. Eine weitere Variable betrifft die Entwicklung der Richtergehälter – ob sie real gestiegen oder gesunken sind und

10 Hierzu und zum Folgenden *Lars P. Feld/Stefan Voigt*, *Economic Growth and Judicial Independence: Cross Country Evidence Using a New Set of Indicators*, *Euro. J. Pol. Econ.* 19 (2003), 497, 500ff.

wie sie im Verhältnis zu Universitätsprofessoren, Anwälten und Ministern stehen – sowie die Ausstattung der Gerichte mit weiteren Mitarbeitern und Ressourcen.

### c) Ergebnisse

Am Ergebnis der *de iure*-Unabhängigkeit überrascht zunächst, dass an der Spitze der Tabelle keiner der OECD-Staaten figuriert und Länder wie die Vereinigten Staaten von Amerika nur im Mittelfeld rangieren. Die Schweiz fällt gar ganz aus dem Rahmen und liegt bei der *de iure*-Unabhängigkeit auf einem der hinteren Plätze (67 von 71). Darin zeigt sich, dass die richterliche Unabhängigkeit nicht plausibel auf eine Analyse der in Verfassung und Gesetz enthaltenen Garantien gestützt werden kann. Bei solchen Garantien verhält es sich manchmal geradezu umgekehrt zur Wirklichkeit: Je weniger ein Land auf die Bekundung der Unabhängigkeit angewiesen ist, desto weniger formale Sicherungen werden sichtbar; je stärker dagegen die Zweifel an der Unabhängigkeit sind, desto mehr fühlen sich Verfassungs- und Gesetzgeber genötigt, die eigenständige Position der dritten Gewalt zu bekunden. Ein Defizit an politischer Kultur soll durch ein Plus an juristisch-formalen Sicherungen kompensiert werden, ohne dass dies im Ergebnis immer gelingt.

Diese Einschätzung wird gestärkt durch ein anderes Ergebnis der Studie. Die gemessene *de facto*-Unabhängigkeit weicht von der *de iure*-Unabhängigkeit so stark ab, dass die „Top Ten“ des jeweiligen Rankings überschneidungsfrei sind. Keiner der *de iure* als besonders unabhängig eingestuften Staaten kann also *de facto* in der Spitzengruppe der richterlichen Unabhängigkeit mithalten. Die nach der *de facto* Analyse entstehende Reihenfolge ist intuitiv plausibler, weil sich am Ende der Liste weder die Schweiz noch die EU-Mitgliedstaaten finden.

Da es sich um eine ökonomische Studie handelte, wurde außerdem nach der Korrelation der Indizes und einzelner Variablen mit der Wirtschaftskraft gefragt. Zwischen *de facto* Unabhängigkeit und Bruttoinlandsprodukt ließ sich ein starker Zusammenhang feststellen, weshalb die Autoren interpretatorisch davon ausgehen, dass die tatsächliche richterliche Unabhängigkeit einen wichtigen Beitrag zum Wirtschaftswachstum leisten kann. Zwischen *de iure* Unabhängigkeit und Wirtschaftskraft bestand hingegen kein Zusammenhang. Lediglich für die Variable, die eine Verfassungsgarantie der Unabhängigkeit beschreibt, ließ sich ein positiver Zusammenhang feststellen. Daraus wurde interpretatorisch gefolgert, dass es ökonomisch betrachtet zwar hilfreich sei, die unabhängige Position des höchsten Gerichts in der Verfassung zu garantieren, dass aber abgesehen von dieser Maßnahme die Absicherung durch rechtliche Garantien allein keine messbare Auswirkung auf die Wirtschaftskraft habe.

#### d) Gegenüberstellung zur Rechtsvergleichung

Die schwierige Frage ist nun: Wie verhalten sich solche empirischen Studien zur Rechtsvergleichung, und was können sie uns nützen?

Dazu ist zunächst festzuhalten, dass diese Studien bereits Rechtsvergleichung *sind*, weil eine Bestandsaufnahme und Gegenüberstellung der Rechtsvorschriften und der Rechtspraxis stattfindet. Allerdings wäre uns diese Art des Vorgehens nicht genug. Bei aller Unbestimmtheit, die es in der Rechtsvergleichung gerade im öffentlichen Recht methodisch gibt, ist klar, dass wir immer bewertend vorgehen.<sup>11</sup> Wir bleiben also bei einem Länderbericht und bei einer Gegenüberstellung nicht im objektiv Messbaren der formalen Garantien stehen, sondern stellen diese in den Zusammenhang mit der Rechtspraxis. Recht untersuchen wir als *angewandtes* Recht.

Übertragen auf die richterliche Unabhängigkeit: Eine Verfassungsgarantie gilt uns nicht schon dadurch als stabil, dass die Verfassung formal nur durch eine qualifizierte Mehrheit geändert werden kann, sondern erst dadurch, dass diese Hürde in der Staatspraxis tatsächlich beschränkend wirkt und nicht etwa durch eine starke Parlamentsmehrheit oder einen dominanten Präsidenten beiseite geschoben wird. Was in den Rechtsdokumenten als Stärke erscheint, kann sich mit Blick auf die Praxis als bedeutungslos und damit schwach erweisen. Umgekehrt würden wir auch die Schwächen einer rechtlichen Regelung nicht unbewertet in den Rechtsvergleich einführen. Hier kann die Schweiz als Beispiel gelten, die formal betrachtet geradezu erschreckende Abberufungsmöglichkeiten für Richterinnen und Richter kennt. Außerdem steht die schweizerische Richterschaft in ständiger Abhängigkeit von periodischer Wiederwahl. Eine Nichtwiederwahl kann sich im Einzelfall als informeller Ersatz für eine Abberufung darstellen. In einer normativen Rechtsvergleichung würden wir diese Faktoren bewerten. Dazu berücksichtigt man dann auch die informellen Kräfte einer Rechtskultur, die im Falle der schweizerischen Justiz moderierend wirken. Bei aller Kritik würde die Schweiz darum in einem normativen Rechtsvergleich mit anderen Ländern nicht außergewöhnlich schlecht abschneiden.

Stellen wir die normative Rechtsvergleichung den empirischen Studien gegenüber, so zeigt sich, dass sie in ihrem bewertenden Vorgehen zwischen den *de iure*- und *de facto*-Indikatoren liegt. Sie ist stärker praxissensitiv als die bewertungsarme empirische Messung der *de iure*-Unabhängigkeit. Gleichzeitig gibt sie – weil sie bloß individuell einschätzt und nicht nach den Regeln der empirischen Sozialforschung misst – die tatsächliche Unabhängigkeit weniger präzise wieder als der *de facto*-Indikator der empirischen Studie.

<sup>11</sup> Vgl. Fabian Wittreck, Die Verwaltung der Dritten Gewalt, 2006, S. 527 ff. m. w. N. in Fn. 2 und 9.

## e) Nutzen für die Rechtsvergleichung?

Wenn schon die normative Rechtsvergleichung mit einem Fuß auf dem Boden der Rechtswirklichkeit steht, hat dann die empirische Forschung noch einen spezifischen Zusatznutzen für die Rechtsvergleichung? Lässt sie sich vielleicht sogar methodisch integrieren?

Hinsichtlich des *de iure*-Index ist der Nutzen offenbar sehr begrenzt. Das empirische Ergebnis kann als Mahnung dienen, wie wenig es auf die formellen Garantien in Rechtsdokumenten ankommt. Aber dieser Mahnung bedarf eine aufgeklärte Rechtsvergleichung gar nicht, ist sie ihr doch ohnehin vertraut.

Nützlicher ist der *de facto*-Index der juristischen Unabhängigkeit. Er geht wirklich über den Erkenntnisstand normativer Rechtsvergleichung hinaus. Vereinfacht gesagt: In juristischer Vergleichung *ahnen* wir nur, inwieweit die Rechtspraxis das Bild der formalen Garantien relativiert; mit den Mitteln der empirischen Sozialforschung *wissen* wir es in einem quantifizierbaren und darüber hinaus intersubjektiv verbindlicheren Sinn. Die anspruchsvollere Analyse der empirischen Forschung kann uns so vor pauschalen Bewertungen bewahren. Sie mag nur selten zu Ergebnissen führen, die wir nicht ohnehin auch intuitiv in unserer Einschätzung der Rechtswirklichkeit erreicht hätten. Aber sie kann zum Beispiel kontrollieren, ob eine in der Tendenz richtige Einschätzung ihrem Ausmaß nach übertrieben gewichtet wurde.

Noch einmal zum Beispiel der richterlichen Unabhängigkeit: Schaut man sich die Praxis in der Schweiz an, so finden sich tatsächlich Fälle einer faktischen Abberufung durch Nichtwiederwahl und auch vereinzelt Beispiele für einen politisch erzwungenen Rücktritt von Richtern. Es wäre aber unangemessen, die Unabhängigkeit in der Schweiz wegen solcher Einzelfälle pauschal als gering einzustufen. Davor kann eine empirische Studie bewahren, wenn sie außer der kritischen Richtung auch das messbare Gewicht dieser Praxis belegt.

Schließlich gibt es durchaus nützliche Überraschungen durch empirische Forschung – also Ergebnisse, die nicht ohne weiteres der Intuition entsprechen, rückblickend aber durchaus plausibel sind. Eine weitere Studie von *Feld* und *Voigt* zeigt beispielsweise, insoweit wenig überraschend, dass eine wirksame Gewaltenteilung, gefördert durch eine ausgebaute Verfassungsgerichtsbarkeit, positiv mit dem Wirtschaftswachstum korreliert und darum ökonomisch erstrebenswert ist. Überraschend ist allerdings, dass dies nicht für die Teilfaktoren der Verfassungsgarantie und des ausgebauten Zugangs zur Verfassungsgerichtsbarkeit gilt. Eine in der Landesverfassung garantierte Normenkontrollbefugnis und ein auf Individualbeschwerden ausgebautes Verfahren nehmen an der grundsätzlich positiven Wirkung des Rechtsschutzes auf die Wirtschaft also nicht teil.<sup>12</sup> Aus ökonomischer Sicht wäre es darum kontra-

12 *Lars P. Feld/Stefan Voigt*, Making Judges Independent – Some Proposals Regarding the Judiciary, in: Roger D. Congleton/Birgitta Swedenborg (Hrsg.), *Democratic Constitutional Design and Public Policy – Analysis and Evidence*, 2006, S. 251–

produktiv, einen so weitgehenden Ausbau des Verfassungsrechtsschutzes anzustreben.

Wir sind frei, aus juristischer Sicht im Ergebnis anderer Meinung zu sein. Einen zusätzlichen Erkenntniswert haben solche Ergebnisse aber allemal, können sie doch in der Rechtsvergleichung etwas genauer verständlich machen, warum längst nicht alle Staaten geradezu automatisch ihre Rechtsschutzinstrumente erweitern.

### 3. Fazit

So viel zur systemunabhängigen empirischen Forschung. Zusammenfassend lässt sich festhalten, dass sie in ihrer bewertungsarmen Bestandsaufnahme formaler Rechtsinstrumente keinen Gewinn gegenüber normativer Rechtsvergleichung darstellt. In der Beurteilung von rechtspraktischen Wirkungen ist sie indes überlegen, weil die Ergebnisse empirischer Sozialforschung intersubjektiv verbindlich, im Ausmaß quantifizierbar und im Einzelfall sogar überraschend sind. Die Ergänzung der Rechtsvergleichung um Empirie empfiehlt sich besonders, wenn man im normativen Rechtsvergleich ein Auseinanderfallen von Rechtsregelung und Rechtspraxis *prima facie* feststellt, sich aber über das Gewicht dieser Abweichung nicht klar ist und darum in der Gesamtbewertung unsicher.

## IV. Systemspezifische empirische Forschung

Die systemspezifische empirische Forschung ist, wie eingangs betont, insofern näher an der Rechtswissenschaft, als sie ihre Fragestellung bereits auf die Institutionen innerhalb des einzelnen Rechtssystems abstellt. Gleichzeitig scheidet damit die Vergleichbarkeit der Ergebnisse zunächst einmal aus. Was beispielsweise zum Präsidentenamt in den Vereinigten Staaten von Amerika erforscht wird, hat für das Amt des französischen Staatspräsidenten keine Bedeutung und umgekehrt.

Trotzdem meine ich, dass sich die systemspezifische empirische Forschung für Rechtsvergleichung fruchtbar machen lässt. Als Beispiel dafür kann der von *André Bächtiger* und anderen empirischen Politikwissenschaftlern entwickelte *Discourse Quality Index* gelten.

### 1. *Discourse Quality Index* (DQI)

#### a) Fragestellung

Die von der Diskurstheorie inspirierte Theorie der *deliberativen Demokratie* weist als Legitimationsgrund des Verfahrens auf das Vorhandensein von realen Diskursen im Recht hin. Das ist in jeder Hinsicht umstritten; einerseits

---

288, zitiert nach der Publikation in SSRN Working Paper Series Nr. 597721 (= CESifo Working Paper Nr. 1260), S. 3, 18 ff., 27.

im theoretischen Ansatz, andererseits bei der Frage, wo *überhaupt* diskursive Elemente in der Staatspraxis auszumachen sind. Die empirische politische Wissenschaft geht davon aus, dass strategische und diskursive Elemente regelmäßig gemeinsam auftreten. Die Fragestellung lautet dann: Welches Gewicht haben die diskursiven Elemente in den jeweiligen Institutionen?

Gemessen werden soll also, wie weit eine bestimmte politische Debatte von dem nie erreichbaren Diskursideal entfernt ist.<sup>13</sup> Lässt sich das für eine Institution bestimmen und auf konkrete Faktoren zurückführen, so könnte es als weiterer Schritt möglich sein, das Design der Institution so zu beeinflussen, dass der deliberative Modus der Entscheidungsfindung gefördert und damit die Legitimationswirkung des Verfahrens gestärkt wird.<sup>14</sup> Und hier kommt die Vergleichung ins Spiel: Die in verschiedenen Staaten praktizierten Verfahrensweisen bilden das Reservoir, aus dem sich Optimierungsvorschläge gewinnen lassen.

## b) Vorgehen

Zur Untersuchung der Diskursqualität wurden einzelne Parlamentsthematen in der Schweiz, den USA und Deutschland betrachtet. Gegenstand waren dabei insgesamt 50 Debatten in verschiedenen parlamentarischen Teilinstitutionen (Kammern, Ausschüssen, Vermittlungsgremien). Diese Debatten wurden einer quantitativen Inhaltsanalyse unterzogen, allerdings nicht als repräsentativer Debattenquerschnitt, sondern im Stil von Fallstudien. Damit scheidet aus empirischer Sicht die Methode einer direkten Vergleichung aus. Die Fragestellung kann nicht mehr als systemunabhängig gelten, sondern beschränkt sich auf die spezifische Situation einer einzelnen Debatte in einer einzelnen Institution (*typological sampling*).<sup>15</sup>

Unter den Themen fanden sich stark polarisierende, beispielsweise aus dem Arbeitsrecht, und weitgehend polarisierungsfreie, beispielsweise zum Tierschutz.

Zur Messung der Diskursqualität wurden sieben Teilindikatoren bestimmt, unter ihnen das störungsfreie Sprechen (*participation*), das Begründen (*justification*) einschließlich der bekundeten Gemeinwohlorientierung (*common good*) und vor allem, als mehrfach unterteilter Indikator, der Respekt (*respect*). Die Debatten wurden dazu in einzelne Sprechakte unterteilt und für jeden Indikator anhand einer Bewertungstabelle mit Diskursivitätswerten versehen. Obwohl dabei sehr viel Kontextberücksichtigung durch die Kodierer nötig war, ließ sich die Kodierungshomogenität dank der einfach gehaltenen Klassifizierungen als sehr gut einstufen.

13 André Bächtiger, *The Real World of Deliberation. A Comparative Study of Its Favorable Conditions in Legislatures, 2005*, S. 29, 33, 36.

14 Bächtiger (Fn. 13), S. 39.

15 Bächtiger (Fn. 13), S. 77 ff., 83 ff.

### c) Ergebnisse

Im Ergebnis der Studie lässt sich Deliberation tatsächlich als eine *graduelle* Eigenschaft in allen untersuchten Debatten nachweisen, sogar in solchen, bei denen es um Verteilungsfragen ging, so dass thematisch eine strategische Verhandlung nahe gelegen hätte. Außerdem zeigt sich, dass der institutionelle Rahmen maßgeblich auf die Art der Debatte einwirkt, also beispielsweise dieselben Parlamentarier in nicht-öffentlichen Kommissionssitzungen anders verhandeln als im öffentlichen Plenum.

Die Einzelanalysen führten allerdings zu einem disparaten Set an Befunden. Nachweislich günstig wirkte es sich auf das Maß an Diskursivität aus, wenn als institutioneller Rahmen eine große Koalition oder eine Vetomacht bestand. Erwartungsgemäß günstig war außerdem, wenn die Debatte zu einem Thema mit geringer Polarisierung geführt wurde.<sup>16</sup> Nicht nachweisbar war der hypothetisch vermutete positive Einfluss einer präsidentiellen Organisationsform, wie sie in den USA existiert. Nur ansatzweise sichtbar wurde die höhere Diskursivität bei der Diskussion in einer zweiten Kammer – das heißt, einer nicht präparativen „Reflexionskammer“ wie dem Senat, dem Bundesrat oder dem Ständerat. In unterschiedliche Richtungen schließlich wirkten die Teilindizes der Diskursqualität bei der nicht-öffentlichen Themenbearbeitung: Einerseits war der Respekt größer, andererseits wurde weniger begründet als im öffentlichen Forum. Verallgemeinbare Ratschläge für ein institutionelles Design waren bei dieser Vielfalt der Befunde bisher kaum möglich.

Trotzdem gibt es einzelne Studienteile, die ein interessantes Licht auf die Systemunterschiede werfen. Nah am Diskursideal bewegte sich beispielsweise eine Debatte des schweizerischen Ständerats – das ist die kleine Kammer mit Kantonsvertretern – zum wenig kontroversen Thema des Sprachenartikels: Hier gab es in öffentlicher Debatte praktisch keine abschätzigen Bemerkungen über den politischen Gegner und der Anteil von Äußerungen mit hohen Respektwerten lag bei fast 70%. Am anderen Ende des Spektrums findet sich eine Debatte des Deutschen Bundestags zum Ladenschlussgesetz: Fast 72% der Einzeläußerungen enthielten respektlose Bemerkungen und die Sprechakte mit hohem Respektwert lagen bei unter 5%. Selbst wenn die Studie keine ursächliche Verknüpfung mit dem institutionellen Rahmen der Parlamentsgremien nachweisen kann, bieten diese Zahlen einen Eindruck von der *Größenordnung des Kulturunterschieds* in den politischen Gremien.

### d) Gegenüberstellung zur Rechtsvergleichung

Stellt man die empirische Herangehensweise der normativen Rechtsvergleichung gegenüber, so fällt auf, dass sie es deutlich schwerer hat, zu Ergebnissen zu gelangen. In der Rechtsvergleichung würde es uns genügen, die verfassungsbedingten Strukturunterschiede der politischen Systeme aufzuzeigen.

<sup>16</sup> *Bächtiger* (Fn. 13), S. 53, 152 ff.

Diese Analyse würden wir dann mit der Rechtsprechung des jeweiligen Verfassungsgerichts und mit unserem persönlichen Eindruck aus der praktischen Funktionsweise der Gremien unterlegen. Als Ergebnis genügt uns häufig ein Panorama der Beschreibungen, manchmal ergänzt durch eine abstrakte Strukturkritik, etwa wenn wir mangelhafte Gewaltenteilung oder Demokratiedefizite aufdecken. Das rechtsvergleichende Vorgehen ist insofern nicht viel anders, als es früher unter dem Titel „Allgemeine Staatslehre“ betrieben wurde.

Die empirische Forschung möchte hingegen das Funktionieren der Institutionen auf die einzelnen Faktoren ihrer Struktur *konkret und nachweisbar* zurückführen. Weder Beschreibung noch prinzipielle Kritik genügen ihr dafür. An die Stelle von Wirkungsvermutungen treten Messungen. Im Idealfall kann die empirische Forschung darum institutionelle Defizite verbindlicher aufdecken und überzeugendere Empfehlungen für ein verbessertes Institutionsdesign abgeben.

#### e) Nutzen für die Rechtsvergleichung?

Gerade weil wir in der normativen Rechtsvergleichung in aller Regel nur die Wirkungsrichtung institutioneller Eigenarten angeben können, aber nichts über das *Ausmaß* ihrer Wirkung im Ensemble der Staatsstruktur wissen, liefert uns die empirische Forschung ergänzende Hinweise. Der Einfluss politischer und sozialer Kultur lässt sich ohne Empirie nur erahnen. So kann man zwar wissen, dass die Konsensorientierung in der schweizerischen Kultur ein stärkeres Gewicht hat als in der deutschen. Man kann den Unterschied auch spüren, wenn man sich als Zuschauer in eine Debatte des Nationalrats setzt. Aber welches intersubjektiv vermittelbare *Gewicht* dieser Kulturbestandteil hat, kann man überzeugend erst nach empirischer Analyse der Debatten festhalten.

Ogleich es sich um systemspezifische empirische Forschung handelt, lassen sich vereinzelt auch konkrete Forderungen daraus in den Rechtsvergleich importieren. So ist es ein schlichter Normbefund, dass trotz verfassungsrechtlich freien Mandats in den deutschen Parlamenten eine Fraktionsdisziplin herrscht, die durch Fraktionsgeschäftsordnungen formalisiert ist. Der deliberative Schaden, der dadurch angerichtet wird, lässt sich ohne Empirie allerdings nicht ermitteln. Die eben vorgestellte Studie zur Diskursqualität hat gezeigt, dass es für die Deliberation förderlich ist, wenn an die Stelle einer polarisierenden Rollenverteilung in Regierungs- und Oppositionsfraktion entweder eine große Koalition oder eine Vetomacht tritt. Es wäre darum zu erwarten, dass die Diskursqualität im Deutschen Bundestag erheblich gesteigert werden könnte, wenn man das Verfassungsrecht des freien Mandats als ein *Verbot der Fraktionsdisziplin* interpretierte.

Die voraussichtlich große Wirkung dieses kleinen Schritts vermag man einzuschätzen, wenn man die normative Betrachtung durch einen empirischen Befund ergänzt. Erst durch eine solche Einschätzung lässt sich begründen, wa-

rum ein Verbot der Fraktionsdisziplin als Ergebnis der Abwägung sinnvoll ist, obgleich auch gewichtige gegenläufige Erwägungen angestellt werden müssen. So wird die Zuordnung der Verantwortlichkeit schwieriger, wenn die Fraktionen nicht mehr im Block über die Politik abstimmen (Repräsentativfunktion). Auch wird es in der Praxis häufiger zu Situationen kommen, in denen langwierige parlamentarische Vorarbeiten letztlich nicht fruchtbar werden, weil eine erwartete Mehrheit sich in der Schlussabstimmung nicht einstellt (Parlamentseffizienz). Trotz solcher Nachteile für die Repräsentativfunktion und die Parlamentseffizienz sollte man allerdings das Verbot der Fraktionsdisziplin, eben weil es nachweislich starke Auswirkungen auf das Diskursniveau hat, als eine insgesamt für die Demokratie förderliche Regelung einstufen.

## 2. Erweiterung der Fragestellung

Man kann die Analyse der Diskursqualität weiter vorantreiben, wenn man nach themenspezifischen Besonderheiten fragt. Schon in der ursprünglichen Studie stellte sich heraus, dass bei den Parlamentsdebatten zum Schwangerschaftsabbruch sowohl in Deutschland als auch in der Schweiz eine relativ zu den USA höhere Diskursqualität erreicht wurde.<sup>17</sup> Dies mag unter anderem daran gelegen haben, dass die amerikanische Debatte auf Polarisierung angelegt war, während man sich in Europa bewusst bemühte, diese bei dem moralisch heiklen Thema zu vermeiden. In Deutschland befreiten die Fraktionen ihre Abgeordneten sogar vom Fraktionszwang. Gegenwärtig erforschen wir in Bern darum, ob sich in der schweizerischen Bundesversammlung bei im engeren Sinne *moralischen* Themen (z.B. Behindertenschutz, Stammzellforschung) ein höheres Diskursniveau feststellen lässt als bei technisch-materiellen Themen (z.B. Steuerpolitik, Marktregulierung).<sup>18</sup> Unter Umständen würde es sich normativ anbieten, hier besondere Verfahrensgefäße zu schaffen.

Eine zweite Entwicklungsrichtung, die wir derzeit in Bern untersuchen, betrifft die Änderung in der politischen Kultur. Allgemein besteht in der Schweiz der Verdacht, dass bedingt durch den Erfolg der rechtsnationalen Schweizerischen Volkspartei (SVP) der politische Umgang im Land rauer geworden sei. Wir versuchen das durch eine Längsschnittstudie zu analysieren, mit der wir zu dem gleich bleibenden Thema der Ausländerpolitik die Diskursqualität der Parlamentsdebatten aus den letzten zwei Jahrzehnten auswerten. Solche Weiterungen der Fragestellung sind allerdings so spezifisch für das einzelne Land, dass sich kein rechtsvergleichender Gewinn mehr einstellen dürfte.

17 Bächtiger (Fn. 13), S. 185.

18 Weitere Informationen unter [www.bids.unibe.ch](http://www.bids.unibe.ch).

### 3. Fazit

Als kurzes Fazit zur systemspezifischen empirischen Forschung lässt sich festhalten, dass sie nur punktuell für die Rechtsvergleichung von Interesse ist. Eigentlich kann sie nur das Gewicht einer rechtsnormativen Besonderheit für ein einzelnes Land begründen. Mittelbar lässt sich aus einer solchen Gewichtung aber unter Umständen ein Argument für die Rechtsvergleichung gewinnen.

## V. Ergebnisse

Wenn ich knapp zusammenfasse, dann kann die empirische Forschung auf mindestens drei unterschiedliche Arten in rechtsvergleichende Studien einbezogen werden. Das *empirische Benchmarking* ist die naheliegendste Variante, die allerdings nur für wenige Rechtsgebiete fruchtbar gemacht werden kann. Die *systemunabhängige empirische Forschung* kann vor allem ein rechtsvergleichend bereits festgestelltes Auseinanderfallen von Rechtsregelung und Rechtspraxis dem Gewicht nach bewerten. In seltenen Fällen führt sie auch zu Erkenntnissen, die nicht ohnehin schon intuitiv durch unseren Blick auf die Rechtspraxis gewonnen werden. Die *systemspezifische empirische Forschung* hat rechtsvergleichend vor allem einen heuristischen Wert: Sie kann auf Faktoren aufmerksam machen, die sonst im Systemvergleich nicht genügend berücksichtigt würden.

Wenn man diese drei Stoßrichtungen der methodischen Einbeziehung trotz ihrer Unterschiedlichkeit gesamthaft betrachtet, dann wirken Empirie und Rechtsvergleichung wie unterschiedliche Waagen. Die Rechtsvergleichung ist so schnell und ungenau wie die mit freier Hand gehaltene Balkenwaage der Justitia. Die sozialwissenschaftlich betriebene Empirie verhält sich demgegenüber wie eine elektronische Präzisionswaage. Häufig müssen wir es in der Rechtsvergleichung gar nicht so genau wissen, aber wo es wirklich auf Details ankommt, ist es gut, auch noch ein präzises empirisches Instrument zu haben.