

---

**Geschäftsführung**

---

**RA Berthold Welling**

VCI e.V.  
Neustädtische Kirchstraße 8  
10117 Berlin  
Tel. (030) 20 05 99 10  
Fax (030) 20 05 99 99  
welling@vci.de

---

**RA Dr. Andreas Richter**

P+P Pöllath + Partners  
Potsdamer Platz 5  
10785 Berlin  
Tel. (030) 253 53 653  
Fax (030) 253 53 800  
berliner.steuergespraech@pplaw.info

---

**Online-Veranstaltung - 77. Berliner Steuergespräch**

**„Evaluierung von Normen – Möglichkeiten und Grenzen“**

Tagungsbericht von *Berthold Welling*  
und *Dr. Andreas Richter LL.M.* <sup>1</sup>

Berlin, den 11. Januar 2021

Im Idealfall sollte Recht einfach, verständlich und zielgerichtet ausgestaltet sein. Ob eine Norm jedoch die Anforderungen hinsichtlich Wirksamkeit, Vollzugstauglichkeit und Adressatenfreundlichkeit tatsächlich erfüllt, zeigt sich erst nach ihrer Umsetzung. Um Verbesserungen einzuarbeiten und den Erfolg einer Rechtsvorschrift bewerten zu können, hat sich der Gesetzgeber vorgenommen, das entstandene Recht systematisch zu evaluieren.

Für eine zweckgerichtete Gesetzgebung kommt es nicht lediglich auf die Effektivität – d. h. die tatsächliche Anwendung des Rechts – an. Zugleich ist von Bedeutung, ob und in welchem Zeitraum die angestrebten Zwecke erreicht wurden, welche Wirkungszusammenhänge und Nebenwirkungen entstanden sind und ob die angefallenen Kosten in einem angemessenen Verhältnis zu den Wirkungen stehen. Seit 2013 soll von dem jeweils federführenden Ressort grundsätzlich bei allen Gesetzen und Verordnungen ab einem jährlichen Erfüllungsaufwand von mehr als 1 Mio. Euro nach fünf Jahren ab Inkrafttreten der Norm eine Evaluierung durchgeführt werden. Die Bundesregierung hat sich in ihrem Arbeitsprogramm „Bessere Rechtsetzung und Bürokratieabbau 2018“ zum Ziel gesetzt, die Vornahme von Evaluationen als Standard zu etablieren. Durch einen Beschluss des Staatssekretärsausschusses „Bürokratieabbau und Bessere Rechtsetzung“ im November 2019 wurde das Evaluierungskonzept der Bundesregierung fortentwickelt und konkretisiert.

Das 77. Berliner Steuergespräch unter der Leitung von Herrn *Prof. Dr. h.c. Rudolf Mellinghoff*<sup>2</sup> bot den Referenten Frau *Prof. Dr. Johanna Hey*<sup>3</sup> und Herrn *Prof. Dr. Thiess Büttner*<sup>4</sup> sowie den Mit-Diskutanten Frau *Dr. Ramona Voshage*<sup>5</sup>, Herr *Felix Hierstetter*<sup>6</sup> und Herrn *Hanns-Eberhard Schleyer*<sup>7</sup> zusammen mit den online Teilnehmenden eine Gelegenheit, um die Bedeutung der Evaluierung von (Steuer-)Normen für eine gute Gesetzgebung und die Wirksamkeit des relativ jungen Evaluierungssystems zu beleuchten und zu diskutieren.

## A. Referate

### I. Evaluierung von Normen – Möglichkeiten und Grenzen

*Prof. Hey* führte mit ihrem Referat allgemein in das Thema der Evaluierung von Normen ein.

#### 1. Begrifflichkeiten und Technik

Ihrer Ansicht nach müssten Gesetze aufgehoben oder geändert werden, wenn sie ihre Ziele verfehlten. Um die Wirksamkeit eines Gesetzes beurteilen zu können, bedürfe es der Evaluierung. Idealerweise solle die Evaluierung schon im Rahmen des Gesetzgebungsverfahrens stattfinden (prospektive Evaluierung). Bei Anwendung dieser Technik bleibe allerdings eine nachträgliche Überprüfung der Regelung weiterhin erforderlich. Das Hauptproblem der prospektiven Evaluierung sei jedoch der zeitliche Aufwand, der mit dieser Methode einhergehe. In Rahmen großer Reformen könne mit Planspielen vorgegangen werden, für die üblichen Gesetzesvorhaben sei dies jedoch zu zeitaufwendig. Daher werde in der Regel retrospektiv evaluiert. Diese Technik sei sicherer, da zu dem Zeitpunkt der Evaluierung bereits verlässliche Daten vorlägen. Für die Glaubwürdigkeit der Ergebnisse sei maßgeblich, ob die Evaluierung intern oder extern durchgeführt worden sei. Die Umsetzung externer, nicht durch Regierung oder Ministerien überwachter Evaluierungen sei daher erforderlich.

Um zu überprüfen, ob das jeweilige Gesetz seine Ziele erreiche, müssten diese Ziele zunächst erkennbar sein. Dazu gebe es in der Geschäftsordnung des Bundestages (GO BT) und in der Gemeinsamen Geschäftsordnung der Bundesministerien (GGO) entsprechende Vorgaben. Trotzdem erfolge häufig eine nicht hinreichend exakte Formulierung der Ziele, so dass diese keine taugliche Grundlage für eine Evaluierung darstellten. Zudem sei das Thema der zeitgerechten Datenverfügbarkeit äußerst relevant. Die Evaluierung solle zum einen nicht allzu spät erfolgen, zum anderen müsse ein gewisser Zeitraum seit der Gesetzgebung verstrichen sein, um auf verlässliche Daten zurückgreifen zu können.

Mit Blick auf die Verfassungsrechtsprechung zur Evaluation müsse eine Abgrenzung zwischen der Evaluierung von Gesetzeswirkungen einerseits und Beobachtungspflichten andererseits erfolgen. Beobachtungspflichten verlangten, dass Anpassungen bei Veränderungen der tatsächlichen und gesetzlichen Rahmenbedingungen der Norm vorzunehmen seien. Dies sei vor allem bei Typisierungen – wie zum Beispiel bei den Einheitswerten – besonders relevant. Bei einer solchen Nachbesserung handle es sich jedoch nicht um eine klassische Evaluierung, die sich auf Gesetzeswirkungen beziehe.

---

<sup>1</sup> *Dr. Andreas Richter, LL.M.* ist Rechtsanwalt, Fachanwalt für Steuerrecht und Partner der Kanzlei POELLATH, Berlin. *Berthold Welling* ist Rechtsanwalt und Geschäftsführer des VCI e.V., Frankfurt am Main, und für Recht, Steuern und Nachhaltigkeit verantwortlich. Beide Autoren sind geschäftsführende Vorstandsmitglieder des Berliner Steuergespräche e.V.

<sup>2</sup> *Prof. Dr. h.c. Rudolf Mellin* ist ehemaliger Präsident des Bundesfinanzhofs, München.

<sup>3</sup> *Prof. Dr. Johanna Hey* ist Leiterin des Instituts für Steuerrecht an der Universität zu Köln.

<sup>4</sup> *Prof. Dr. Thiess Büttner* ist Inhaber des Lehrstuhls für Volkswirtschaftslehre, insbesondere Finanzwissenschaft, an der Friedrich-Alexander-Universität Erlangen-Nürnberg.

<sup>5</sup> *Dr. Ramona Voshage* ist Leiterin der Abteilung „Gesamtwirtschaft“ im Amt für Statistik Berlin-Brandenburg, Potsdam.

<sup>6</sup> *Felix Hierstetter* ist Steuerberater und Leiter des Bereichs „Steuern International“ bei der General Electric Deutschland Holding GmbH, Frankfurt am Main.

<sup>7</sup> *Hanns-Eberhard Schleyer* ist Mitglied des Nationalen Normenkontrollrats, Berlin.

## 2. Evaluierungsbedarf im Steuerrecht – Unterscheidung nach Normtypen

Da der Gesetzgeber sehr viele Normen schaffe, müsse ausgewählt werden, welche Normen einer Evaluierung unterzogen werden müssten. In diesem Zusammenhang werde im Steuerrecht zwischen verschiedenen Normtypen unterschieden. Bei Fiskalzwecknormen erfolge in Gesetzesbegründungen eine Schätzung ihrer Haushaltsauswirkungen in Form von Steuer mehr- oder -mindereinnahmen. Es sei wichtig, im Nachhinein zu evaluieren, ob diese Annahmen auch tatsächlich eingetreten seien. Die geschätzten Steuer mehr- und -mindereinnahmen spielten in der Politik eine große Rolle, da sie Gesetze scheinbar messbar machten. Klar sei auch, dass Lenkungsnormen evaluiert werden müssten. So hänge zum Beispiel die Rechtfertigung von Steuervergünstigungen von deren Wirksamkeit ab. Daher sei zu erwarten, dass bei diesen Normen stets überprüft werde, ob das Lenkungsziel eingetreten sei. Auch Vereinfachungsnormen müssten in Hinblick auf drei Aspekte einer Evaluierung unterzogen werden. Zunächst müsse geprüft werden, ob die Typisierung realitätsgerecht sei. Darüber hinaus bestehe eine Beobachtungsnotwendigkeit bzw. das Erfordernis der Fortschreibung. Zudem müsse überprüft werden, ob die bezweckten Vereinfachungseffekte bzw. Bürokratie-Einspareffekte tatsächlich eingetreten seien. Bei Missbrauchsvermeidungsnormen schließlich sei es wünschenswert, die missbräuchliche Verhaltensweise sowie das Ausmaß der missbrauchsbedingten Steuerausfälle und Auswirkungen im Vorhinein offenzulegen, da nur dann die Wirksamkeit der Missbrauchsbekämpfungsregelung beurteilt werden könne.

## 3. Gesetzesevaluierung steuerrechtlicher Normen – Status quo

Bei der Evaluierung von Bürokratiekosten handle es sich um die Hauptaufgabe des Normenkontrollrats (vgl. § 2 Absatz 3 Gesetz zur Einsetzung eines Nationalen Normenkontrollrates - NKRG). Daneben führten auch die Rechnungshöfe Evaluierungen durch, welche sie für kritische Stellungnahmen zu bestehenden Regelungen und die Vorstellung entsprechender Reformvorschläge nutzten. Ein Beispiel dafür sei der Bericht des Bundesrechnungshofes nach § 88 Absatz 2 der Bundeshaushaltsordnung (BHO) zum Entwurf des Forschungszulagengesetzes (FZulG). Jedoch verfügten die Rechnungshofberichte über keine unmittelbaren rechtlichen Wirkungen. Auch das Bundesfinanzministerium (BMF) evaluiere grundsätzlich intern, gelegentlich komme es aber auch zur Vergabe externer Evaluierungsaufträge. Ein Beispiel dafür sei die Überprüfung von 33 Steuervergünstigungen in den Jahren 2017 bis 2018 durch ifo Institut, Fifo Köln, ZEW und Fraunhofer-Institut.

## 4. Institutionalisierung der Gesetzesevaluierung

Das Hauptproblem der Evaluierung liege darin, dass politisches Evaluationsversprechen und gesetzliche Evaluationspflicht in einem Spannungsverhältnis stünden. Besonders bei politisch umstrittenen Vorhaben verspreche der Gesetzgeber häufig die Durchführung der Evaluation, die am Ende jedoch ausbleibe. Ein Beispiel hierfür sei die Zinsschranke – diese sei bis heute nicht evaluiert worden.

Zunächst stelle sich die Frage, ob es überhaupt rechtliche Evaluierungspflichten gebe. In den §§ 43 Absatz 1 Nr. 5, 44 Absatz 7 GGO gebe es die Gesetzesbegründung betreffende formelle Vorgaben. Danach müsse in der Gesetzesbegründung auch zu den Gesetzesfolgen Stellung genommen werden. Zudem müsse bereits in der Gesetzesbegründung festgelegt werden, ob und

in welchen Zeiträumen die Evaluation durchgeführt werden solle. Auch in Gesetzen gebe es Evaluierungsklauseln, beispielsweise in § 33b Absatz 8 Einkommensteuergesetz (EStG), § 17 FZulG und § 23 Bundesemissionshandelsgesetz (BEHG).

Es stelle sich die Folgefrage, welche anderen Möglichkeiten es gebe, um die Wirksamkeit von Gesetzen rechtlich zu sichern. In § 111 GGO NRW sei beispielsweise eine Befristungspflicht für Gesetze vorgesehen. Auf diese Weise werde der Gesetzgeber nach Ablauf der Befristung dazu angehalten, die Wirksamkeit der Norm zu überprüfen. Jedoch sei die Effektivität dieser Methode fraglich, da es häufig zu einer automatischen Verlängerung befristeter Normen komme. Aufgrund des Erfordernisses der hinreichenden Gesetzesbestimmtheit könne die Evaluierung zudem niemals eine Bedingung für die Fortgeltung einer gesetzlichen Regelung sein.

## 5. Verfassungsrechtliche Relevanz

Folglich sei festzuhalten, dass es sowohl einfach-rechtliche Instrumente als auch gesetzliche Evaluationsklauseln gebe, deren Durchsetzbarkeit jedoch fraglich seien. Daher sei die Frage aufzuwerfen, ob das Erfordernis der Evaluierung verfassungsrechtliche Relevanz habe. Es müsse unterschieden werden, ob überhaupt eine prozedurale Evaluierungspflicht bzw. Beobachtungspflicht bestehe, und was folge, sobald die Unwirksamkeit einer Norm festgestellt worden sei. Strittig sei, ob sich aus der Verfassung Anforderungen an das innere Gesetzgebungsverfahren ableiten ließen. Man könne nach Willi Geiger den Standpunkt vertreten, dass „der Gesetzgeber nichts als das Gesetz schulde“, so dass keine Begründungspflicht bestehe. Nach Ansicht des Bundesverfassungsgerichtes (BVerfG) sei die Kehrseite des Prognosespielraums des Gesetzgebers, der im Rahmen des Gesetzgebungsverfahrens großzügige Annahmen und Typisierungen ermögliche, dass die jeweilige Ausgestaltung im Nachhinein überprüft werden müsse. Durch das Verfassungsgericht werde auch gefordert, dass der Gesetzgeber Fehlannahmen korrigieren müsse. Zudem existierten viele Entscheidungen zu den Beobachtungspflichten bei veränderten Rahmenbedingungen, zum Beispiel in den Fällen W-Besoldung für Professoren, Hartz IV und § 6a EStG.

Im Falle der W-Besoldung entwickelte das BVerfG eine prozedurale, verfassungsrechtliche Anforderung an Begründung, Überprüfung und Beobachtung der Besoldung. Die Verfassungswidrigkeit solle bereits dann vorliegen, wenn den vom Gericht aufgestellten Anforderungen nicht genügt werde. Auch in der Hartz IV-Entscheidung sei für die Verfassungskonformität der Hartz IV-Sätze ein sachgerechtes und transparentes Verfahren gefordert worden, welches eine realitätsgerechte und nachvollziehbare Ermittlung der Hartz IV-Beträge ermögliche. In beiden Entscheidungen habe das BVerfG Anforderungen an das Gesetzgebungsverfahren mit verfassungsrechtlichen Folgen bei Verletzung der Voraussetzungen gestellt. Im Zusammenhang mit § 6a EStG habe das Verfassungsgericht entschieden, dass der Gesetzgeber bei einer wesentlichen Änderung der Rahmenbedingungen überprüfen müsse, ob Anpassungsbedarf bestehe. Der Forderung nach realitätsgerechter Typisierung sei immanent, dass die tatsächlichen Begebenheiten bei der Normschaffung zugrunde gelegt werden müssten, woraus eine verfassungsrechtliche Evaluationspflicht hergeleitet werden könne.

Fraglich sei, ob bereits das Unterlassen der Evaluierung oder erst das Vorliegen einer verfehlten bzw. nicht realitätsgerecht typisierten Norm verfassungsrechtlich beachtlich sei. Sofern sich die Norm aufgrund der Verfehlung der gesetzgeberischen Ziele als unwirksam erweise, könne dies auch Konsequenzen für die Verfassungsmäßigkeit der Regelung nach sich ziehen. Grundsätzlich

müsse die Unwirksamkeit einer Norm zu ihrer Verfassungswidrigkeit führen. Die Schwierigkeit liege jedoch in der Festlegung, wann eine Norm derart ihre Ziele verfehle, dass die Annahme der Verfassungswidrigkeit angemessen sei. Dies solle nur bei Extremfällen angenommen werden. In diesem Zusammenhang müsse der weite Beurteilungs- und Prognosespielraum berücksichtigt werden, den das BVerfG dem Gesetzgeber einräume.

## **6. Problem: Rechtsschutz gegen unwirksame Normen**

Gegen belastende Normen, die ihre gesetzlichen Ziele verfehlten, könne der Steuerpflichtige Rechtsschutz suchen. Im Falle einer Art negativen Klage gegen ineffiziente Steuervergünstigungen, die (möglicherweise) zu nicht erforderlichen Mitnahmeeffekten führten, werde jedoch mangels Beschwer und Rechtsschutzbedürfnisses kein Rechtsschutz gewährt. Die Erfolgsaussichten einer solchen Klage seien zudem im Hinblick auf den Beurteilungsspielraum des Gesetzgebers gering. Allenfalls bei offensichtlicher Zweckverfehlung, zum Beispiel bei einem strukturellen Erhebungsdefizit, bestehe Aussicht auf Erfolg eines Rechtsmittels.

## **7. Fazit**

Die Evaluierung von Steuergesetzen sei zwar wünschenswert und notwendig, aber auch bei gesetzlicher Regelung kaum rechtlich durchsetzbar. Es bestehe auch keine allgemeine verfassungsrechtliche Pflicht zur Evaluierung, an die der Gesetzgeber gebunden sei. Er könne lediglich mit externem Druck dazu angehalten werden, Evaluierungen durchzuführen und Normen nachträglich anzupassen.

## **II. Zielsetzung und Möglichkeiten der Evaluierung steuerlicher Normen**

*Prof. Büttner* befasste sich in seinem einführenden Referat mit der Zielsetzung und den Möglichkeiten der Evaluierung steuerlicher Normen.

Laut dem Bericht „Notwendigkeit, Potenzial und Ansatzpunkte einer Verbesserung der Dateninfrastruktur für die Steuerpolitik“ (vom 30.10.2020) des Wissenschaftlichen Beirats des BMF existiere ein Spannungsverhältnis: Zum einen gebe es „über die Wirkungen des deutschen Steuersystems und seiner spezifischen Regelungen auf zentrale wirtschaftliche und soziale Zielgrößen in vielen Bereichen nur Mutmaßungen“, zum anderen seien „in der quantitativen Analyse der Besteuerung in den vergangenen Jahrzehnten weltweit erhebliche Fortschritte gemacht worden“. Das Gutachten lege ein Maßnahmenpaket für eine durchgreifende und nachhaltige Verbesserung der Dateninfrastruktur für die Steuerpolitik vor, da die derzeitigen Maßnahmen für eine systematische Evaluierung nicht ausreichten.

### **1. Zwecke der quantitativen Analyse steuerlicher Normen**

Die quantitative Analyse habe drei Zwecke: Der erste sei die Evaluierung des Gesetzgebungsprozesses. Der Gesetzgeber sei auf Informationen über die Wirkungsweise der Steuergesetze angewiesen, um auf deren Grundlage qualifizierte Entschlüsse fassen zu können. Zweitens seien der Steuervollzug und die Vollzugsdefizite von Relevanz. Maßgeblich sei, wie die Steuergesetze in der Praxis bei der Interaktion zwischen Steuerpflichtigem und Finanzbehörde wirkten. Dabei

sei bekannt, dass viele Vollzugsdefizite existierten. Zuletzt sei die quantitative Analyse Grundvoraussetzung dafür, dass eine (verfassungs-)rechtliche Kontrolle von Steuergesetzen erfolgen könne.

## 2. Evaluierung

Im Rahmen einer idealen Steuerpolitik werde im Rahmen eines Evaluierungskreislaufes regelmäßig und systematisch evaluiert. Die Evaluierung solle dafür im ersten Schritt ex-ante, also vor Beschließung des Gesetzes, durchgeführt werden. Im Vorfeld könnten bereits diverse Erwartungen festgehalten werden: die Ziele der Norm, Art und Weise der Ermittlung der Zielerreichung, der „Zielerreichungsgrad“, die Höhe des erwarteten Steueraufkommens, der Träger der Last der zu beschließenden Steuer sowie die Bürokratie- und Folgekosten. Nach der Implementierung könne im zweiten Schritt anhand einer ex-post-Betrachtung beurteilt werden, ob die Erwartungswerte erreicht oder verfehlt worden seien, und welche Gründe es für etwaige Abweichungen gebe. Auf dieser Grundlage könne diskutiert und neue Initiativen entwickelt werden, um die gesetzlichen Parameter zu verbessern. Dies führe wiederum zur Schaffung neuer Normen, so dass der Evaluierungskreislauf erneut beginne. In der Praxis existiere kein derart ideales Evaluierungssystem. Daher sei in der Steuerpolitik ein inkrementelles Vorgehen erforderlich.

## 3. Aufkommenseffekte

In Deutschland schätze der Gesetzgeber die Steuereinnahmen zweimal jährlich für das laufende Jahr voraus. Bei dieser Steuerschätzung handle es sich um eine koordinierte Vorausschätzung der Einnahmen für die Haushaltsaufstellung aller staatlichen Ebenen. Die Vorausschätzung ist maßgeblich für die Entwicklung eines Haushaltsgesetzes. Da Fiskalregeln strukturellen Ausgleich erforderten, sei eine verlässliche Vorausschätzung zu den tatsächlichen Ausgaben erforderlich. Die Aufstellung einer verlässlichen Vorausschätzung sei schwierig. Vorhersagefehler resultierten vor allem aus der – nicht verlässlich vorherzusehenden – konjunkturellen Entwicklung sowie aus Aufkommenseffekten von Steuerreformen. Da die Werte bei der Berechnung bereinigt seien, führten schwankende Einnahmen allein nicht unmittelbar zu der Erforderlichkeit einer Haushaltsanpassung.

Bei Steuerreformen sei dies nicht so einfach. Erfahrungen aus der Praxis zeigten deutlich, dass es zu einer systematischen Überschätzung der Aufkommenseffekte von Steuererhöhungen und -senkungen komme. Eine Steuersatzerhöhung habe keine proportionale Steigerung der Einnahmen zur Folge, sondern einen Aufkommenseffekt, der niedriger liege. Grund dafür sei, dass Akteure, die von den steuerlichen Maßnahmen betroffen seien, versuchten, belastete steuerliche Aktivitäten herunterzufahren und stattdessen unbelastete auszuweiten. Die Diskrepanz zwischen dem prognostizierten und dem tatsächlichen Aufkommen resultiere aus der Reaktion der Bemessungsgrundlage, die sich nicht linear fortschreiben lasse, sondern durch Verhaltensreaktionen beeinflusst werde. Aus dem ex-post-Schätzwert lasse sich schließen, dass der Schätzfehler zwischen  $\frac{1}{4}$  und  $\frac{1}{3}$  des Mehraufkommens liege. Die Reaktion der Bemessungsgrundlage sei für die Bewertung des Steuersystems von großer Bedeutung, denn sie falle für jede Steuer unterschiedlich aus und sei ein Maß dafür, wie teuer und effektiv die Steuererhebung sei.

Idealerweise müsse eine optimale Steuerpolitik durch den Einsatz verschiedener Behandlungs- und Kontrollgruppen erprobt und ermittelt werden. Ein solch experimentelles Vorgehen sei jedoch

nicht möglich. Die Angemessenheit einer Steuerpolitik müsse daher anhand anderer Kriterien bewertet werden, zum Beispiel durch die Auswertung früherer Reformen sowie durch das Aufdecken von Sprungstellen (wie zum Beispiel Höchstabzugsbeträge, Mindestbeträge etc.). In den letzten Jahren habe es erhebliche methodologische Fortschritte bei datenintensiven Verfahren gegeben (zum Beispiel Paneldatenanalyse, Diskontinuitätsanalyse, „Matching“-Verfahren, Instrumentvariablen, synthetische Kontrollgruppen). Folglich gebe es viele Möglichkeiten, Evaluierungen durchzuführen. Mit Hilfe der Mikrodaten der Steuerzahler könnten kausale Effekte identifiziert und eine gesetzliche Regelung quantitativ evaluiert werden.

#### **4. Steuerlücke**

Die Evaluation sei auch im Zusammenhang mit Steuerlücken – sogenannten Tax gaps – von großer Bedeutung. Eine Steuerlücke sei der Unterschied zwischen erwarteten und tatsächlichen Einnahmen einer Steuer. Einige Staaten erhoben in regelmäßigen Abständen Steuern zur Deckung solcher Steuerlücken. Zwar erfolge in Deutschland keine systematische Ermittlung von Steuerlücken, jedoch seien diese oft Motivation für Reformen und Maßnahmen – Beispiele dafür seien die Unternehmenssteuerreform 2008 sowie die Einfügung des § 35a EStG zur Bekämpfung der Umsatzsteuerhinterziehung durch Schwarzarbeit. Evaluation könne sowohl zur Ermittlung von Steuerlücken als auch zur Schaffung von Gegenmaßnahmen hilfreich sein.

#### **5. Potenziale administrativer Daten**

Im steuerlichen Bereich seien die verfügbaren Daten repräsentativ, valide sowie aktuell und bildeten somit eine verlässliche Grundlage für Analysen. Aufgrund der fortgeschrittenen Digitalisierung der Steuerverwaltung seien die Kosten für die Datenerhebung und Informationssammlung vergleichsweise gering. Der Schutz der Daten sei durch Anonymisierung gewährleistet, durch welche jedoch das Analysepotenzial nicht verloren gehe. Die Durchführung einer gründlichen und systematischen Evaluierung sei kompliziert und aufwendig. Daher sei es von Vorteil, dass Wissenschaftler weltweit die Daten unabhängig und unentgeltlich auswerten. Für eine erfolgreiche Durchführung sei jedoch eine Evaluierungskultur erforderlich. Dafür müsse eine große Zahl an Wissenschaftlern mit den Daten vertraut sein, die Methoden der Evaluation korrekt anwenden können und zur Verbesserung der Techniken in ständigem Austausch stehen.

Die Ausgangslage in Deutschland sei in diesem Zusammenhang nicht optimal. Gesetzliche Vorgaben für die Nutzung der Steuerdaten seien zu eng und bestehende Nutzungsmöglichkeiten der Mikrodaten seien unterentwickelt. Auch im bestehenden Datenangebot gebe es erhebliche Defizite, da die Steuerstatistik viele Steuern nicht erfasse. Im Vergleich dazu sei die Datenlage bei den Parafisci – zum Beispiel dem Institut für Arbeitsmarkt- und Berufsforschung (IAB) oder der Rentenversicherung – deutlich besser. Die Daten des IAB bildeten daher die Grundlage für die Evaluierung der Arbeitsmarktpolitik.

#### **6. Empfehlungen**

Der Bericht des Wissenschaftlichen Beirats des BMF empfehle vor allem drei Maßnahmen. Zum einen solle ein eigenes Forschungsdatenzentrum für Steuern eingerichtet und Wissenschaftlern

der Zugang zu den Daten ermöglicht werden, zum anderen sollten Statistiken erweitert und verbessert werden sowie Verknüpfungsmöglichkeiten zwischen Statistiken geschaffen bzw. ausgeweitet werden.

Im Anschluss an die einführenden Referate von Prof. Hey und Prof. Büttner erteilte Prof. Mellinghoff den weiteren Mit-Diskutanten das Wort für eine erste Stellungnahme.

### **III. Die Perspektive der amtlichen Statistik**

*Dr. Voshage* habe die Potenziale amtlicher und administrativer Daten früh erkannt. Daher befürworte sie im Rahmen des gesetzlichen Auftrages unter Wahrung des Datenschutzes eine Zurverfügungstellung der Daten. Dies sei die Grundlage für externe Evaluierungen. Es treffe zu, dass die bestehenden Nutzungsmöglichkeiten der Mikrodaten unterentwickelt seien. Die derzeit vorhandene Daten-Infrastruktur müsse dringend modernisiert werden. Die Statistischen Ämter in Deutschland seien zur Datenerhebung mandatiert und sammelten viele Daten zu verschiedenen Themen wie Wirtschaft, Gesellschaft, Gesundheit und Bildung. Dazu zähle auch die Steuerstatistik, die sich unter anderem mit Einkommen-, Gewerbe- und Umsatzsteuer befasse. Durch die Datenerhebung, die auf gesetzlicher Grundlage erfolge, finde ein Eingriff in die informationelle Selbstbestimmung der Bürger statt. Daher müssten – vor allem bei der Erfüllung der Auskunftspflichten durch die Ämter – der Datenschutz und die Geheimhaltung gewährleistet werden. Die Statistischen Ämter verfügten über eine große Zahl an Datensätzen, welche unter Beachtung der Datenschutzstandards über die Forschungsdatenzentren der Wissenschaft zur Verfügung gestellt werden sollten. Dieser Datenzugang sei die Grundlage für eine externe Evaluierung durch die unabhängige Wissenschaft. Durch die Einrichtung der Forschungsdatenzentren werde zugleich eine organisatorische, rechtliche und technische Infrastruktureinrichtung geschaffen. Insgesamt gebe es 150 verschiedene Statistiken mit über 1000 Datenprodukten, die von der Wissenschaft ausgewertet werden könnten. Durch technische und methodologische Weiterentwicklungen sei die Nachfrage nach Daten in den letzten Jahren rasant gestiegen. Daher sei eine evidenzbasierte Politikberatung inzwischen einfacher umsetzbar. Die erforderliche Daten-Infrastruktur, welche auch für die Ministerien große Bedeutung habe, sei vorhanden.

### **IV. Die Sicht des Nationalen Normenkontrollrats**

Herr *Schleyer* stellte in seinem Beitrag den Nationalen Normenkontrollrat sowie dessen Aufgaben und Initiativen vor. Der Normenkontrollrat existiere seit 2006 und agiere als unabhängiges Beratungsgremium der Bundesregierung. Das Gremium finde seine gesetzliche Grundlage im NKRG und bestehe aus 10 Mitgliedern, mit einem Frauenanteil von 70%. Zu den Aufgaben des Normenkontrollrats gehöre vor allem die Prüfung der Kostenfolgen von Normen für Bürger, Wirtschaft und öffentliche Verwaltung. Zudem erarbeite er Stellungnahmen zu alternativen Gesetzesvorlagen oder schlage die Konzeption von Alternativvorschlägen vor.

Darüber hinaus habe sich der Normenkontrollrat stets als Gremium verstanden, welches einen Beitrag zur besseren und transparenteren Rechtssetzung leisten wolle. Daher sei der Kreis der Aktivitäten in den letzten Jahren erweitert worden. Ein Beispiel dafür seien die „Einfacher zu...“-Projekte. Im Mittelpunkt dieser Projekte stehe die Untersuchung des Vollzugs von Bundesgesetz-



zen über die verschiedenen Ebenen von Bund, Ländern und Kommunen hinweg auf Möglichkeiten der Vereinfachung und Beschleunigung. Im Rahmen dieser Projekte sei beispielsweise thematisiert worden, wie der Bezug eines BAföG-Darlehens für Studierende oder der Erhalt einer Arbeits- und Aufenthaltserlaubnis für ausländische Arbeitnehmer vereinfacht werden könne. Durch diese Projekte seien Verfahren erfolgreich vereinheitlicht und beschleunigt sowie gemeinsame Empfehlungen verschiedener Vereinigungen und Verbände erarbeitet worden.

Der Normenkontrollrat habe sich in den letzten Jahren auch darum bemüht, neue Instrumente zu schaffen, um bestehende Probleme zu lösen. Beispielsweise sei festgestellt worden, dass die EU bei der Vorbereitung und Schaffung von Normen häufig die Kostenfolgen für Wirtschaft, Bürger und Verwaltungen der Mitgliedstaaten nicht hinreichend berücksichtigt habe. Mit dem EU Ex-ante-Verfahren, an dessen Durchsetzung der Normenkontrollrat wesentlich mitgewirkt habe, überprüfe die Bundesregierung die Folgekosten von EU-Rechtsakten für Deutschland erstmals systematisch. Nach dem „One in, one out“-Verfahren des Normenkontrollrats müsse eine Belastung der Wirtschaft in Höhe eines Betrages X an einer anderen Stelle durch eine Einsparung in derselben Höhe wieder ausgeglichen werden. Mit Hilfe der Bundesregierung sei es gelungen, die Einführung des Verfahrens durchsetzen, dessen Umsetzung vom Normenkontrollrat streng kontrolliert werde. Im nächsten Schritt solle dieses Verfahren auf die europäische Ebene übertragen werden; Bundeskanzlerin Angela Merkel habe sich zustimmend zu diesem Vorhaben geäußert.

Allgemein sei bessere Gesetzgebung nur durch den Einsatz von Evaluierung möglich. In 24 von 43 OECD-Staaten gebe es bereits Evaluierungsstandards und Evaluierungshilfen. Der Normenkontrollrat plädiere dafür, in Deutschland ein vergleichbares System einzuführen. Der erste Versuch im Jahr 2013 sei aus Sicht des Normenkontrollrats nicht zielführend gewesen, da es dem Konzept an Qualitätsstandards und Qualitätssicherung gefehlt habe. Vor allem dank der Bemühungen und Vorschläge des Normenkontrollrats habe der Staatssekretärs-Ausschuss der Bundesregierung im November 2019 ein umfassendes Evaluierungskonzept beschlossen und auf diese Weise das bestehende Konzept fortentwickelt und konkretisiert. Der Beschluss führe zwar lediglich zu einer Selbstbindung der Regierung, bewirke aber trotzdem eine erhebliche qualitative Verbesserung des Evaluierungsprozesses. Die wichtigste Neuerung sei die Etablierung von Anforderungen an Evaluierungsregelungen: Erstens müsse aus diesen klar werden, welche Ziele ein Gesetz verfolge, zweitens, welche Kriterien für die Beurteilung der Zielerreichung zugrunde gelegt worden seien und drittens, auf welcher Datengrundlage diese Ziele erreicht werden sollen.

Nach Ansicht des Staatssekretärs-Ausschusses müsse die Bundesregierung nicht jedes Vorhaben evaluieren, welches zu einer Belastung von mehr als 1 Mio. Euro führe. Im Umkehrschluss bedeute dies aber, dass zumindest alle wichtigen Normen evaluiert werden müssten. Es gehöre nun zum Zuständigkeitsbereich des Normenkontrollrats, die ordnungsgemäße Umsetzung dieses Evaluierungskonzepts zu überprüfen. Bereits bei der Erarbeitung von Referentenentwürfen achte der Normenkontrollrat darauf, dass eine Evaluierungsklausel vorgesehen sei, welche den Anforderungen dieses Konzepts genüge. Evaluierungsklauseln könnten entweder im Gesetzestext verankert oder in die Gesetzesbegründung eingebaut werden.

Aus dem Beschluss des Staatssekretärs-Ausschusses folgten zwei weitere Aufgaben für den Normenkontrollrat, die bislang noch nicht erfüllt worden seien: Zum einen müsse die Entwicklung einer Arbeitshilfe zu einer Standardisierung führen; zum anderen müsse eine zentrale Einheit geschaffen werden, welche die Schaffung der Rahmenbedingungen für die Durchführung der

Evaluierung überprüfe, den Evaluierungsprozess koordiniere und bei Bedarf entsprechende Handlungsempfehlungen konzipiere.

## V. Evaluierung aus der Sicht der Wirtschaft

Nach Ansicht von Herrn *Hierstetter* sei das Thema Evaluierung nicht nur für die Verwaltung, sondern auch für Unternehmen – insbesondere für solche, die grenzüberschreitend tätig seien – äußerst wichtig. Evaluierung werde häufig im Zusammenhang mit Aufkommenseffekten erwähnt. In den vergangenen Jahren sei zunehmend in den Fokus der Wirtschaft gerückt, wie handhabbar Steuerrechtsnormen und wie groß Rechtssicherheit sowie administrativer Aufwand bei der Anwendung dieser Normen seien. Da in letzter Zeit mehrere Entscheidungen zu diesem Thema gefällt worden seien, sei es auch zu einer Verkomplizierung der Rechtslage gekommen. Da das nationale Recht zunehmend durch europäische Rechtssetzung beeinflusst werde, spielten vor allem die europäischen Richtlinien in diesem Zusammenhang eine entscheidende Rolle. Daher müsse bei der Evaluierung an den richtigen Stellen – nämlich auf Ebene der EU – angesetzt werden.

EU-Rechtssetzung wirke sich nicht nur in Deutschland, sondern in allen Mitgliedstaaten und im Falle einiger Rechtsakte sogar außerhalb der EU aus. Die deutsche Wirtschaft sei international vernetzt und verfüge in vielen EU-Mitgliedstaaten über Standorte. Die europäischen Rechtsakte der jüngsten Zeit verursachten eine hohe Belastung und viel Komplexität. Dies könne anhand der Richtlinie für die Verwaltungszusammenarbeit (DAC, Directive on Administrative Cooperation) veranschaulicht werden. Kürzlich sei DAC 6 in Kraft getreten. Dadurch seien unter anderem „Country-by-Country-Reporting“ und Anzeigepflichten bei grenzüberschreitenden Transaktionen relevant geworden. Die DAC-Richtlinie führe zu einer permanenten Belastung der Unternehmen, vor allem durch Systemumstellungen – insbesondere im Reporting-Bereich – und durch Pflichten zur Informationsbereitstellung. Die Unternehmen bräuchten Zeit, um sich mit den neuen Verfahren und Verpflichtungen vertraut zu machen.

Ein weiteres aktuelles Thema sei ATAD (Anti Tax Avoidance Directive), die Anti-Steuervermeidungsrichtlinie der EU. Es komme zunehmend zu Verknüpfungen zwischen ausländischen Tatbestandsmerkmalen und inländischen Rechtsfolgen. Auch daraus resultierten Rechtsunsicherheit, Komplexität und ein erhöhter administrativer Aufwand. Beispielhaft seien „Hybrid mismatch“- , „Subject to tax“- sowie „Controlled foreign corporation-rules“ (CFC) erwähnt. Dabei seien neue Rechtsbegriffe oft nicht hinreichend klar definiert oder es gebe Überschneidungen und Wechselwirkungen, welche die Rechtsanwendung erschwerten. Zudem sei weder bei ATAD noch bei der DAC-Richtlinie eine vollständige Harmonisierung erfolgt. Die Rechtsanwendung werde sowohl für die Unternehmen als auch für die Verwaltung dadurch noch komplexer, dass die Richtlinien lediglich Mindeststandards etablierten, von denen die Mitgliedstaaten bei der Ausgestaltung abweichen könnten. Die EU schaffe vermehrt Recht, welches den Unternehmen immer mehr Reporting-Pflichten auferlege, im Gegenzug komme es aber nicht zur Schaffung entlastender Mechanismen.

Insgesamt könnten die entsprechenden Richtlinienregelungen dem Anspruch an Gesetze hinsichtlich der Zielgerichtetheit, Einfachheit und Anwendungsfreundlichkeit nicht gerecht werden. Die Kommission habe bereits vor einigen Jahren Richtlinien zur Evaluierung verabschiedet. In diesen sei der Ablauf einer Evaluierung der EU-Richtlinien relativ detailliert geregelt. Vorgesehen seien eine ex ante-Folgenabschätzung, Konsultationen und ein ex post-Monitoring-Verfahren. Vor

einigen Jahren habe die Bundesregierung in Zusammenarbeit mit dem Normenkontrollrat ein ex ante-Verfahren eingeführt, dessen Zweck es sei, den Prozess auf europäischer Ebene so früh wie möglich durch nationale Maßnahmen zu begleiten. Eine nationale Prüfung, die erst dann erfolge, wenn das europäische Gesetzgebungsverfahren bereits weit fortgeschritten sei, sei nicht effektiv. Das „One in, one out“-Verfahren werde auf europäischer Ebene noch nicht zur Anwendung gebracht, so dass die Bürokratiebremse aktuell nur auf nationaler Ebene Wirkung entfalte.

Die ex ante-Folgenabschätzungen bzw. Evaluierungen der EU seien oftmals zu oberflächlich und nicht belastbar. So sei beispielsweise das Impact assessment zu DAC 6 unzureichend. Im Jahr 2019 sei die DAC-Richtlinie durch die Kommission ex post evaluiert worden. Dabei sei festgestellt worden, dass noch viel Verbesserungsbedarf bestehe. Zum Beispiel sei den Mitgliedstaaten oft nicht bewusst, welche Vorteile ihnen aus den übermittelten Daten entstünden, und wie sie diese nutzen könnten. In der Folge komme es nicht zu einer effektiven Anwendung der Normen, während bereits die baldige Implementierung von DAC 7 angekündigt werde. Evaluierungen seien aber nur dann wirksam, wenn anschließend eine Anpassung der ermittelten Schwachpunkte erfolge.

Im Bereich des Steuerrechts habe das ex ante-Verfahren noch keine relevanten Auswirkungen auf die Praxis gehabt. Es sei wichtig, die Durchsetzung des Verfahrens zu unterstützen. In diesem Zusammenhang werde darauf hingewiesen, dass die Verbände zum Beispiel im Falle von DAC 6 und dem ATAD-Umsetzungsgesetz einerseits spät in das Gesetzgebungsverfahren einbezogen worden seien, andererseits äußerst kurze Stellungnahme-Fristen – bei DAC 6 und ATAD waren es lediglich drei Werkzeuge – eine konstruktive Beurteilung aus Sicht der Unternehmen vereitelten. Die Stellungnahme-Fristen seien in Anbetracht der Komplexität und Relevanz der Regelungen unangemessen kurz gewesen. Eine derart unzureichende Beteiligung der Wirtschaft schmälere die Aussagekraft und Stichhaltigkeit einer ex ante-Folgenabschätzung. Zudem sei fraglich, ob überhaupt die Bereitschaft bestehe, eine europäische Norm nach Durchführung der ex post-Evaluierung umfassend abzuändern.

Daher müsse bereits im Entstehungsstadium von Normen besonders auf Handhabbarkeit und Zielgerichtetheit geachtet sowie eine realistische und konkrete Folgenabschätzung durchgeführt werden. Dafür müssten Experten und die von der Norm Betroffenen, wie zum Beispiel Verbände, frühzeitig und ernsthaft am Gesetzgebungsverfahren beteiligt werden. Die Evaluierungsergebnisse müssten neue Maßnahmen und Veränderungen anstoßen können. Befristete Gesetze seien zwar auf nationaler Ebene nicht zielführend, könnten sich aber auf europäischer Ebene als wirksam erweisen, wenn die Mitgliedstaaten einige Jahre nach Implementierung erneut über eine Norm abstimmen und der Beibehaltung zustimmen müssten. Wenn die ganze EU, und nicht nur ein Nationalstaat allein, eine Norm überprüfe, werde diese einer kritischeren Prüfung unterzogen.

## **B. Diskussion**

*Prof. Mellinshoff* eröffnete die Diskussion mit der Frage an Herrn *Schleyer*, ob der Gesetzgeber nicht aus Prinzip an Gesetzgebungsvorhaben festhalte, auch wenn nach externen Evaluierungen Kritik im Zusammenhang mit dem Vollzug und der Effizienz der Norm vorgebracht worden seien. Es stelle sich die Frage der tatsächlichen Wirksamkeit solcher ex ante-Evaluierungen.

Herr *Schleyer* erwiderte, dass einer der zentralen Forderungen des Normenkontrollrats die tatsächliche Berücksichtigung von Evaluationsergebnissen und die daraus folgende Novellierung von Normen sei. Eine Gesetzesänderung solle grundsätzlich nicht ohne eine umfassende Evaluierung vorgenommen werden. Nach Ansicht des Normenkontrollrats müsse das Gesetzgebungsverfahren modernisiert werden, da es seit 200 Jahren im Wesentlichen nicht verändert worden sei. Die meisten Gesetze beruhten auf Referentenentwürfen. Es gebe kein festgeschriebenes Verfahren, nach dem Interessenverbände in die Vorbereitung des Entwurfes eingebunden werden müssten. Ob und wie der Einbindung stehe im Ermessen jedes Referenten. Daher sei die Berücksichtigung von Fachwissen und Stellungnahmen von Seiten der Wirtschaft in einem Referentenentwurf einzelfallabhängig. Die GGO sehe einen relativ späten Einbezug von Verbänden und Kommunen in das Gesetzgebungsverfahren vor. Die Folgen davon äußerten sich im Rahmen des Vollzugs der Normen durch die Kommunen, da die auf Bundesebene erarbeiteten Vorgaben in der Praxis häufig nicht oder nur schwer umsetzbar seien. Auch die von Herr *Hierstetter* angebrachte Kritik an den kurzen Stellungnahme-Fristen sei berechtigt, da diese der Durchführung einer angemessenen Vorevaluierung im Wege stünden. Die Gesetzgebungsverfahren der letzten Jahre seien oft mit kurzen Fristen und ohne Einbeziehung relevanter Parteien durchgeführt worden – wobei damit nicht die Maßnahmen zur Bekämpfung der Corona-Pandemie gemeint seien.

Die Verschlechterung des Gesetzgebungsprozesses sei ein Standortnachteil für Deutschland. Daher sei ein Paradigmenwechsel in der Gesetzgebung erforderlich. Der Normenkontrollrat habe 2019 ein Gutachten mit dem Titel „Erst der Inhalt, dann die Paragraphen“ veröffentlicht. Darin werde ein neues Vorgehensmodell für die Gesetzesvorbereitung vorgeschlagen. Danach solle die Bundesregierung im ersten Schritt politische Ziele unter Belassung von genügend Spielräumen für die praktische Umsetzung festlegen. Bei der Vorbereitung müsse bereits die Umsetzung des Vorhabens durchdacht und geplant werden, wobei besonderes Augenmerk auf den digitalen Vollzug von Gesetzen gelegt werden müsse. In diesem Stadium müsse bereits eine umfangreiche Konsultation mit den Betroffenen – beispielsweise Verbände, Gewerkschaften und Kommunen – stattfinden. Die Beteiligten erstellten im zweiten Schritt ein gemeinsames Eckwertepapier, welches im Anschluss dem Kabinett vorgelegt werde. Erst nach Zustimmung durch das Kabinett werde der Referentenentwurf auf Grundlage des Eckwertepapiers erstellt. Das maßgebliche Problem des aktuellen Gesetzgebungsverfahrens sei, dass ein bereits erstellter Referentenentwurf in der Regel nicht mehr grundlegend geändert werde. Ein neues Vorgehen bei der Vorbereitung von Gesetzen löse dieses Problem und sei für die Bürger, die Wirtschaft und die öffentliche Verwaltung von Vorteil.

*Prof. Mellinshoff* merkte an, dass aus verfassungsrechtlicher Sicht einerseits fraglich sei, ob die verfassungsrechtliche Grenze der Zumutbarkeit überschritten sein könne, wobei es sich dabei um ein Thema handle, zu dem sich das Verfassungsgericht bisher nicht ausdrücklich geäußert habe. Andererseits könne eine Evaluierungspflicht ggf. über eine entsprechende Änderung der Verfassung verpflichtend in das Gesetzgebungsverfahren implementiert werden.

*Prof. Hey* antwortete, dass eine Stellungnahme des BVerfG wünschenswert sei. Die von Herrn *Hierstetter* angesprochenen Probleme könne man tatsächlich durch eine Grundgesetzänderung lösen. Beispielsweise handle es sich bei den Meldepflichten, welche den Unternehmen auferlegt worden seien, um Eingriffe in ihre Freiheitsrechte. Unter dem Gesichtspunkt der Verhältnismäßigkeit müsse daher zunächst geprüft werden, ob diese überhaupt erforderlich seien.

Zwar führe das BVerfG mit Verweis auf den weiten Gestaltungsspielraum des Gesetzgebers selbst keine Evaluierungen durch. Dennoch sei es dem Verfassungsgericht möglich, die für die Unternehmen steigenden Verwaltungslasten und Meldepflichten für unverhältnismäßig zu erklären. Bei EU-Normen bestehe jedoch das Problem der fehlenden Zuständigkeit des BVerfG.

*Prof. Hey* schloss sich der Ansicht an, dass bei der Gesetzgebung auf EU-Ebene die Höhe der Verwaltungslast oft nicht angemessen berücksichtigt werde. Der von Herrn *Hierstetter* gemachte Vorschlag der Befristung von Normen auf europäischer Ebene sei sinnvoll. Es bestehe aber die Gefahr, dass einer im Grundgesetz festgehaltenen Evaluierungsverpflichtung pro forma oder nur oberflächlich nachgekommen werde, vor allem in Anbetracht der hohen Zahl an Steuergesetzen, bei denen die festgelegte Schwelle von 1 Mio. Euro sehr schnell erreicht werde. Es sei wünschenswert, dass das BVerfG die Höhe von Verwaltungslasten genauer untersuche oder durch geeignete Instrumente eine Verbesserung der Lage erreicht werde. Auch politischer Druck von außen sei hilfreich, weshalb besonders die externe Evaluierung von Bedeutung sei.

Herr *Schleyer* betonte, dass gerade aus diesem Grund eine Qualitätsstandardisierung und -sicherung erforderlich sei. Es müsse klar festgelegt werden, welche Voraussetzungen eine Evaluierung erfüllen müsse. In diesem Zusammenhang sei anzumerken, dass es aus Sicht des Normenkontrollrats keine Unterschiede zwischen einer guten internen und einer guten externen Evaluierung gebe.

*Prof. Mellingshoff* fragte *Dr. Voshage*, ob die Daten aus den 150 erwähnten Statistiken des Statistischen Amtes selbst ermittelt worden seien oder ob die Wissenschaft an die Ämter herantrete und den Zugriff auf bestimmte Datensätze erbitte. Außerdem sei interessant, ob die Finanzverwaltung den Forschungsdatenzentren vorgeben könne, welche Statistiken erstellt werden sollten. Die kleinen Anfragen im Bundestag zeigten, dass die Bundesregierung oftmals nicht über relevante Daten oder Erkenntnisse verfüge, obwohl es die Aufgabe des Staates sei, für die Gesetzgebung relevante Daten rechtzeitig zu erheben und für die Verfügbarkeit bestimmter Rohdaten, auf welche auch die Wissenschaft zugreifen können müsse, zu sorgen.

*Dr. Voshage* erläuterte zunächst die Entstehung der Steuerstatistiken. Mit den Steuergesetzen entstünden auch die Gesetze für die Datenerhebung und Steuerstatistiken. Eine Datenerhebung erfolge nur bei Vorliegen eines entsprechenden Gesetzes. Für die gesetzliche Grundlage sei das Bundesinnenministerium zuständig, da die Statistischen Ämter diesem untergeordnet seien. Um vor allem externe Evaluierungen durch die Wissenschaft zu ermöglichen, müssten die Statistischen Ämter kraft Gesetzes zum Tätigwerden, also zum Erheben weiterer Daten, ermächtigt werden. Die aktuelle gesetzliche Grundlage sei nicht optimal, da zum Beispiel Bayern in diesem Zusammenhang eine abweichende Rechtsauffassung vertrete, und daher die Zurverfügungstellung bayerischer Einzeldaten nur unter erschwerten Bedingungen möglich sei. Wenn mehr Evaluierung gewünscht sei, müsse sich etwas an der Rechtslage ändern. Zudem müsse in den Gesetzen klargestellt werden, wie der Datenzugang geregelt werden solle. Das Forschungsdatenzentrum verfüge über die einzelnen Mikrodaten, aber nicht über aggregierte Daten. Diese stellten die Ämter nur der Wissenschaft zur Verfügung.

*Prof. Büttner* führte aus, dass aus der Diskussion das Erfordernis einer grundsätzlichen und weitreichenden Prüfung und Überarbeitung des Steuersystems deutlich geworden sei. Genauso

müsse auch das Evaluierungssystem geprüft und durchdacht werden. Welche Daten den Statistischen Ämtern zur Verfügung stünden, hänge stark von der Ausgestaltung des jeweiligen Statistikgesetzes ab, wobei es möglich sei, die Statistikgesetze auszuweiten. Bereits jetzt verfügten die Finanzverwaltungen über umfangreiche Daten, zum Beispiele aufgrund der eBilanz. Zudem seien sie verpflichtet, die Daten für weitere Untersuchungen an das BMF zu übermitteln. Im Allgemeinen genüge es nicht, ausschließlich die von den Statistischen Ämtern auf Grundlage der Statistikgesetze erhobenen Daten zu betrachten. Es müssten im Wege einer Gesamtschau auch die unternehmensbezogenen Daten der Finanzverwaltungen berücksichtigt werden. Im Rahmen der Statistikgesetze sei eine derartige Verknüpfung der verschiedenen Daten nicht möglich.

In Deutschland gebe es viele Informationen, die zwar vorhanden seien, aber erst noch verfügbar gemacht werden müssten. Anhand dieser Informationen könnten viele relevante Probleme gelöst werden. Bezüglich der effizienten Nutzung und Zurverfügungstellung von Daten gebe es „Best practices“ aus anderen Staaten. Daher habe der Wissenschaftliche Beirat des BMF die Schaffung eines neuen Steuerdatenzentrums vorgeschlagen, welches die Daten sammle und zentral zur Verfügung stelle. *Prof. Büttner* sehe – anders als Herr *Schleyer* – gravierende Unterschiede zwischen internen und externen Evaluierungen. Interne Evaluierungen seien kaum brauchbar. Zwar könnten intern Rechenkalkulationen durchgeführt werden, aber kritischen Aspekten, wie zum Beispiel Befolgungskosten, Ausweichmöglichkeiten und tatsächlicher Wirksamkeit von Regelungen könne lediglich mit methodischen Annahmen begegnet werden. Diese Annahmen müssten jedoch offen diskutiert werden, was bei internen Evaluierungen nicht möglich sei, so dass das Evaluierungsergebnis nicht belastbar sei.

Nach Ansicht der Leibniz-Institute, welche oft Evaluierungen durchführten, gelte die interne Evaluation der Ministerien, also die Ressortforschung, nicht als echte wissenschaftliche Forschung. Dafür gebe es mehrere Gründe. Zum einen werde diese Forschung von einem bestimmten Auftraggeber bezahlt, zum anderen seien die Ergebnisse oft nicht transparent, da externen Wissenschaftlern der Zugriff auf die verwendeten Daten verwehrt werde. Zur Schaffung von Transparenz müssten die relevanten Daten in anonymisierter Form veröffentlicht und interessierten Wissenschaftlern zugänglich gemacht werden. Dies sei die Grundlage für die Entstehung einer Evaluationskultur. Erst wenn dieser Schritt getan sei, könne man sich Gedanken um die Schaffung geeigneter Rahmenbedingungen für die Evaluierung machen. Der Wissenschaftliche Beirat des Bundeswirtschaftsministeriums habe vor einigen Jahren ein Positionspapier zur evidenz-basierten Politik veröffentlicht, welches Leitlinien für eine belastbare Gesetzesevaluierung enthalte. Auch die von Herr *Hierstetter* erwähnten Compliance-Kosten seien für die Evaluierung bedeutsam. Mit wissenschaftlichen Methoden könnten die Compliance-Kosten genau errechnet werden. Es sei zwar nicht bekannt, wie hoch die Compliance-Kosten des deutschen Steuersystems aktuell seien, aber vermutlich handle es sich um einen erheblichen Betrag.

*Prof. Mellinshoff* verwies darauf, dass ein Identifikationsnummerngesetz in Planung sei, bei dem die Steueridentifikationsnummer mit 200 öffentlichen Registern verbunden werde. Anschließend erfolge eine anonymisierte Veröffentlichung der Daten. Nach Ansicht einiger Datenschutz-Rechtler könne man die Daten jedoch nicht so weit anonymisieren, dass keinerlei Rückschlüsse auf den Steuerpflichtigen oder das Unternehmen gezogen werden könnten.

An Herrn *Hierstetter* richtete er die Frage, inwieweit die Unternehmen vor diesem Hintergrund bereit seien, ihre Rohdaten zur Verfügung zu stellen, um auf Grundlage von amtlichen und betrieblichen Daten eine effektive Evaluierung zu ermöglichen.

Herr *Hierstetter* erwiderte, dass besonders große Unternehmen auch über Ländergrenzen hinweg ohnehin bereits eine Vielzahl von Daten bereitstellen müssten. Daher müsse es auch möglich sein, die Daten an andere Gremien und Institutionen für die Durchführung von Folgenabschätzungen zu übermitteln. Da die Vorteile einer wirksamen Evaluierung groß seien, sei auch davon auszugehen, dass die Unternehmen zu einer Übermittlung der Daten bereit seien, sofern dies Evaluierungszwecken diene. Im Einzelnen hänge die Bereitschaft davon ab, welche Daten betroffen seien und wie ihre Geheimhaltung gewährleistet werde. Zwar sei aus Sicht der Wirtschaft die Wichtigkeit von Aufkommenseffekten nicht von der Hand zu weisen, jedoch werde diese Thematik zunehmend von der Frage überlagert, ob die einschlägigen Steuernormen noch handhabbar seien. Die Probleme, welche beispielsweise aus unbestimmten Rechtsbegriffen, der notwendigen Berücksichtigung ausländischer Rechtsanschauungen für die Anwendung deutscher Steuervorschriften sowie den Vollzugsdefiziten resultierten, könnten nur durch klare Regelungen gelöst werden. Zum Beispiel sei das BMF beim Referentenentwurf des § 4k EStG um eine klare und möglichst kurze Regelung bemüht gewesen. Dennoch werde sich die Anwendung der Norm für Wirtschaft, Verwaltung und Gerichte als schwierig erweisen. Solche Probleme könnten auch ohne die Zurverfügungstellung von Daten gelöst werden.

Die Teilnehmerin Frau *Corina Kögel*<sup>7</sup> stellte fest, dass viel über das Gesetzgebungsverfahren und dessen Qualität gesprochen werde. Die Gesetzgebung obliege dem Souverän, dem Parlament. Vorbereiter der Gesetze seien die Ministerien, im Steuerrecht das BMF. Die technische Bewältigung obliege somit der Exekutive. Wie Herr *Schleyer* bestätigt habe, werde im Gesetzgebungsverfahren nicht viel geändert. Fraglich sei, ob an dieser Stelle nicht etwas grundsätzlich geändert werden müsse, und ob es nicht eines eigenen Apparats bedürfe, um die Entwürfe zu evaluieren.

*Prof. Mellinshoff* sei gegen die Einführung eines separaten „Evaluierungsministeriums“. Stattdessen solle die Evaluierung durch möglichst viele externe Wissenschaftler ermöglicht werden. Nur wenn Evaluierungen verschiedener Wissenschaftler und Institutionen vorlägen, könne die Notwendigkeit und Erforderlichkeit von Normen in sinnvoller Weise diskutiert werden.

Unter Bezugnahme auf *Prof. Büttner*s Beitrag stellte Herr *Schleyer* klar, dass lediglich gemäß der Erfahrungen des Normenkontrollrats die Fremdevaluationen in der Regel qualitativ nicht besser gewesen seien als interne Evaluierungen. Grund dafür seien die finanziellen Mittel, die in den Ressorts für diese Zwecke zur Verfügung stünden. Die ständige Durchführung von Evaluationen sei mit nicht unerheblichen Kosten verbunden. Daher sei es so wichtig, einen ressortübergreifenden Qualitätsstandard für Evaluierungen zu etablieren. Ebenfalls müsse die Qualitätssicherung gewährleistet sein – vor allem bezüglich der internen Evaluierungen.

Abschließend wünschte sich Herr *Hierstetter*, dass der Gesetzgeber die Forderungen des Normenkontrollrats ernst nehme und vor allem das ex ante-Verfahren ordnungsgemäß durchführe, um auf diese Weise einer europäischen Rechtssetzung ohne ausreichende nationale Prüfung Einhalt gebieten zu können.

---

<sup>7</sup> *Corina Kögel* ist Richterin und Vizepräsidentin des Finanzgerichts Hamburg.

Herr *Schleyer* betonte, dass in diesem Zusammenhang viele positive Entwicklungen, wie zum Beispiel die Einführung des EU Ex-ante-Verfahrens oder des „One in, one out“-Verfahrens, zu verzeichnen seien. Er wünsche sich die Umsetzung des neuen Vorgehensmodells für die Gesetzesvorbereitung, wie es in dem Gutachten „Erst der Inhalt, dann die Paragraphen“ vorgesehen sei. Im Bundesjustizministerium fänden aktuell zwei entsprechende Pilotprojekte statt. Die Umsetzung des neuen Modells verspreche eine bessere Rechtsetzung.

*Dr. Voshage* wünsche sich mehr politisches Engagement auf Seiten der Wissenschaftler, um Änderungen der gesetzlichen Grundlagen zu erwirken. Auf diese Weise könne der Zugang zu relevanten Daten ermöglicht und die empirische Arbeit verbessert werden. Derzeit sei der Datenzugang unterentwickelt und müsse modernisiert werden. Hingegen sei die Dateninfrastruktur bereits vorhanden, sie müsse lediglich genutzt werden. Zudem schlug sie vor, die relevanten Diskussionspunkte zum Thema Evaluierung in das Eckpunktepapier der Bundesregierung zur Datenstrategie einfließen zu lassen. Daneben könne erwogen werden, parallel zum Bundesdatenschutzbeauftragten die Position des Bundesdatenbeauftragten, der für das Datenmanagement verantwortlich sei, zu schaffen.

*Prof. Büttner* sei optimistisch, dass die Diskussion zum Thema Evaluierung an den relevanten Stellen weitergeführt werde. Das große Steueraufkommen in Deutschland müsse sinnvoll verwendet werden. Daher müsse transparent sein, wo die Steuerlasten entstünden und wie Steuergesetze funktionierten. Dies wiederum sei nur durch die Zurverfügungstellung von Daten an die Wissenschaft möglich. Es müsse ein Dialog mit der öffentlichen Hand darüber geführt werden, wie die steuerliche Lage verbessert und vor allem Folgekosten abgebaut werden könnten.

*Prof. Hey* befürworte, dass das Thema Evaluierung in der öffentlichen Diskussion mehr Beachtung finde und es zu einer Verbesserung der Evaluierungsstrukturen komme. Es sei traurig, dass bei der Evaluierung einerseits viele Fortschritte gemacht worden seien, Evaluierungen aber andererseits stetig schlechter ausfielen. Jedoch liege das Problem nicht in der Evaluierung, sondern im Ablauf des Gesetzgebungsprozesses. Im Steuerrecht gebe es diverse reformbedürftige Bereiche. Diese müssten einzeln evaluiert und angepasst werden.

### **C. 78. Berliner Steuergespräch**

*Prof. Mellinshoff* dankte allen Diskussionsteilnehmern und schloss die Diskussion mit Hinweis auf die Online-Veranstaltung - 78. Berliner Steuergespräch, das am 19. April 2021 stattfinden wird ([www.berlinersteuergespraech.de](http://www.berlinersteuergespraech.de)).