

ifo BERICHT

Entwicklung eines Bottom-up Ansatzes zur Messung der Unternehmensteuerlücke (Corporate Income Tax Gap) in Deutschland mit Fokus auf die benötigten Datengrundlagen

Kurzexpertise im Auftrag des Bundesministeriums der Finanzen im
Rahmen des Forschungsauftrags fe 3/19

vorgelegt von: *ifo Forschungsgruppe Steuer- und Finanzpolitik
David Gstrein, Elena Herold und Florian Neumeier*

Entwicklung eines Bottom-up Ansatzes zur Messung der Unternehmensteuerlücke (Corporate Income Tax Gap) in Deutschland mit Fokus auf die benötigten Datengrundlagen

Kurzexpertise im Auftrag des Bundesministeriums der Finanzen

*Autor*innen*

David Gstrein

Elena Herold

Dr. Florian Neumeier

Mai 2022

ifo INSTITUT

Leibniz-Institut für Wirtschaftsforschung
an der Universität München e.V.

Forschungsgruppe Steuer- und Finanzpolitik

Auftraggeber:

Bundesministerium der Finanzen, Wilhelmstraße 97, D-10117 Berlin

Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie die Übersetzung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung reproduziert oder unter Verwendung elektronischer Systeme gespeichert, verarbeitet, vervielfältigt oder verbreitet werden.

Inhaltsverzeichnis

1	Einleitung	1
2	Die Unternehmensteuerlücke	3
3	Verfahren zur Schätzung der Unternehmensteuerlücke	5
3.1	Regressionsbasierter Ansatz.....	5
3.1.1	Vorgehen	5
3.1.2	Forschungsergebnisse.....	8
3.1.3	Bewertung.....	10
3.2	Matching-Ansatz.....	11
3.2.1	Vorgehen	11
3.2.2	Forschungsergebnisse.....	12
3.2.3	Bewertung.....	12
3.3	Randomisierte Betriebsprüfungen.....	12
3.3.1	Vorgehen	12
3.3.2	Ergebnisse	14
3.3.3	Bewertung.....	17
3.4	Risikobasierte Betriebsprüfungen.....	18
3.4.1	Vorgehen	18
3.4.2	Ergebnisse	19
3.4.3	Bewertung.....	21
4	Perspektiven für die Steuerverwaltungspraxis	23
5	Zusammenfassung und Schlussfolgerungen	25
	Literaturverzeichnis	27

1 Einleitung

Jährlich entgehen den Staaten weltweit Steuereinnahmen, da die eingekommenen Unternehmensteuern hinter den geschuldeten Unternehmensteuern zurückbleiben. Schätzungen ergeben je nach Land eine Diskrepanz zwischen 4% und 20%. Die Gründe für die Existenz dieser sogenannten Unternehmensteuerlücke sind vielfältig: Ursache kann unbeabsichtigtes Fehlverhalten der Steuerpflichtigen sein, welches sich auf unklare Gesetzgebung und fehlende Kenntnisse zurückführen lässt; kriminelles Verhalten in Form von Steuerhinterziehung; legale Formen der Steuervermeidung und -gestaltung ermöglicht z.B. durch Lücken im Steuergesetz; oder Insolvenz der Steuerpflichtigen (Europäische Kommission, 2018). Der daraus resultierende Verlust an Steueraufkommen erschwert nicht nur das Erreichen fiskalpolitischer Ziele, sondern entzieht dem Staat auch finanzielle Mittel für Investitionen und Umverteilungsmaßnahmen.

Langfristiges Ziel sollte es deshalb sein, die Unternehmensteuerlücke – soweit dies möglich ist – zu verringern. Dadurch ließe sich nicht nur der fiskalische Spielraum des Staates ausweiten, sondern auch die wahrgenommene Steuerfairness und Steuertransparenz erhöhen. Voraussetzung dafür ist allerdings, dass das Ausmaß der entgangenen Steuereinnahmen und die Gründe für den Steuerausfall identifiziert werden. Existierende Verfahren für die Schätzung der Unternehmensteuerlücke lassen sich grob in zwei Gruppen einteilen: Top-down Ansätze (Analysen basierend auf Makrodaten) und Bottom-up Ansätze (Mikro-Analysen auf Basis von Informationen über einzelne Steuerpflichtige).

Die Schätzverfahren variieren stark im Hinblick auf die Daten- und Informationserfordernisse sowie darauf, als wie valide die gewonnenen Ergebnisse einzuschätzen sind. Eine wachsende Anzahl an Ländern veröffentlicht Schätzergebnisse für die Unternehmensteuerlücke, aber auch für die Umsatz- und Einkommenssteuerlücke (OECD, 2017; Europäische Kommission, 2016). Im Rahmen der internationalen Kooperation sowie zur Verbesserung der Vergleichbarkeit der Ergebnisse über Ländergrenzen hinweg ist eine Harmonisierung der Schätzverfahren erstrebenswert. Laut Europäischer Kommission (2018) werden in den europäischen Mitgliedsländern für die Schätzung der Unternehmensteuerlücke bisher vor allem Bottom-up Ansätze verwendet. Als Datenquelle dienen dabei insbesondere die Ergebnisse von Betriebsprüfungen. Existierende Schätzungen konzentrieren sich dabei sowohl auf das gesamte Ausmaß der Steuerlücke, als auch auf deren Ursachen. Die Nutzung von Informationen aus Betriebsprüfungen hat den Vorteil, dass Erkenntnisse über einzelne Steuersegmente und Sektoren, Eigenschaften der geprüften Unternehmen und die Gründe für das Auftreten eines Mehrergebnisses gewonnen werden können. Diese Erkenntnisse können außerdem genutzt werden, um das Risikomanagement im Rahmen der Betriebsprüfung zu verbessern und die Steuerverwaltung effizienter zu machen.

Ziel dieser Kurzexpertise ist es, einen Überblick über Bottom-up Ansätze für die Schätzung der Unternehmensteuerlücke zu bieten. Der Fokus liegt dabei auf solchen Verfahren, die länderüber-

greifend eingesetzt und harmonisiert werden können. Die Expertise kommt dabei zu dem Ergebnis, dass Informationen gewonnen aus randomisierten Betriebsprüfungen als Goldstandard für die Schätzung der Unternehmensteuerlücke anzusehen sind. Darüber hinaus stellt die Kurzexpertise dar, wie die aus Betriebsprüfungen gewonnen Informationen genutzt werden können, um die Unternehmensteuerlücke zu verringern und das Mehrergebnis zu maximieren.

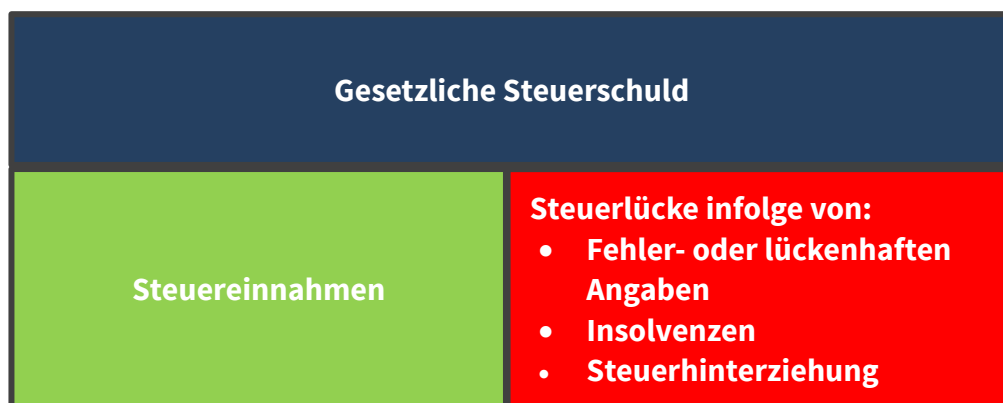
Die Kurzexpertise ist wie folgt aufgebaut. Abschnitt 2 bietet eine kurze Einführung in das Konzept der Unternehmensteuerlücke und erläutert Ursachen für ihre Existenz. In Abschnitt 3 wird beschrieben, welche Bottom-up Verfahren zur Schätzung der Steuerlücke derzeit verbreitet sind. Ferner werden die Anwendungsvoraussetzungen sowie die Vor- und Nachteile der einzelnen Verfahren diskutiert. Abschnitt 4 zeigt Möglichkeiten auf, wie die Praxis der Betriebsprüfungen in Deutschland effizienter gestaltet und das Mehrergebnis maximiert werden kann. Abschnitt 5 fasst die wesentlichen Ergebnisse zusammen.

2 Die Unternehmensteuerlücke

Die Steuerlücke ist ein nützlicher Indikator für die Höhe des jährlichen Steuerausfalls sowie des Steuereinnahmepotenzials. Darüber hinaus kann sie Aufschluss geben, ob es im Rahmen der Steuererhebung und des Steuervollzugs Verbesserungsbedarfe gibt. Für die Schätzung der Unternehmensteuerlücke und die Identifikation ihrer Ursachen gibt es verschiedene Ansätze. Die Hauptproblematik liegt darin, das Steueraufkommen für das kontrafaktische Szenario einer vollkommenen Steuerehrlichkeit zu berechnen. Zusätzlich lassen sich nicht alle Formen von Steuervermeidung und Steuerhinterziehung erfassen, da die betreffenden Unternehmen diese Aktivitäten zu verschleiern versuchen. Schätzungen der Unternehmensteuerlücke sollten daher als das interpretiert werden, was sie sind, nämlich *Schätzungen*. Dabei können die Validität und Reliabilität der Ergebnisse stark vom verwendeten Schätzverfahren und den zugrundeliegenden Daten abhängen.

In einer engen Definition beschreibt die Unternehmensteuerlücke die Differenz zwischen der gesetzlich geschuldeten Steuer und den tatsächlich eingenommenen Steuern (Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2020). Abbildung 1 zeigt die Ursachen für das Auftreten einer Steuerlücke gemäß der engen Definition. Zur Höhe der Steuerlücke können bewusstes Handeln der Steuerpflichtigen, insbesondere in Form von Steuerhinterziehung, Insolvenzen und/oder unbewusste Fehler auf Seiten der Steuerpflichtigen beitragen.

Abbildung 1: Bestandteile der Unternehmensteuerlücke im engeren Sinn



Quelle: Eigene Darstellung

Unbewusstes Handeln/Fehler: Zu einer Diskrepanz zwischen den geschuldeten Steuern und den tatsächlich abgeführten Steuern kann es beispielsweise kommen, weil Steuerpflichtige bei der Steuererklärung unbeabsichtigt Fehler machen oder steuerpflichtige Sachverhalte nicht als

solche identifizieren. Der Beitrag unbewussten Handelns zur Steuerlücke kann verringert werden, indem die Kommunikation der Steuerbehörden verbessert und komplizierte Steuerverfahren vereinfacht werden.

Steuerhinterziehung/Steuerbetrug: Steuerhinterziehung umfasst sowohl die verminderte Angabe also auch die komplette Verbergung von Teilen des zu versteuernden Einkommens und ist illegal. Auf Basis von Top-down Verfahren lässt sich das Ausmaß an Steuerhinterziehung kaum ermitteln. Bottom-up Ansätze wie beispielsweise Schätzverfahren, die auf Informationen aus Betriebsprüfung aufbauen, erfassen Steuerhinterziehung nur, wenn die ökonomischen Aktivitäten auch registriert sind und nicht im informellen Sektor bzw. in der Schattenwirtschaft stattfinden.

Insolvenzen: Auch der Konjunkturzyklus beeinflusst die Steuerlücke. Rezessionen erhöhen das Auftreten von Insolvenzen und tragen damit dazu bei, dass Unternehmen ihre Steuerschuld nicht oder nicht in voller Höhe begleichen können.

In einer weiten Definition wird zur Unternehmensteuerlücke zusätzlich jener Verlust an Steueraufkommen hinzugerechnet, der durch legale Aktivitäten zur Reduzierung der Steuerlast zustande kommt (Europäische Kommission, 2018). In vielen Fällen – insbesondere bei grenzüberschreitenden Aktivitäten – haben Unternehmen rechtliche Gestaltungsspielräume, die es ihnen erlauben, die Steuerlast auf legalem Wege zu reduzieren. In jüngerer Vergangenheit sind beispielsweise Inkonsistenzen zwischen nationalen Steuervorschriften sowie Lücken in Doppelbesteuerungsabkommen als wichtige Ursache dafür identifiziert worden, dass multinationale Konzerne sich der Besteuerung ihrer Gewinne zu einem gewichtigen Teil entziehen konnten. Die Identifikation des Ausmaßes an legaler Steuervermeidung auf Basis von Top-down Ansätzen gestaltet sich durch das Hohe Aggregationsniveau von Makrodaten oftmals als schwierig. Voraussetzung für die Identifikation auf Basis von Bottom-up Ansätzen sind umfangreiche Informationen zu den realwirtschaftlichen Aktivitäten der Unternehmen im In- und Ausland (vgl. Abschnitt 3).

Länder unterscheiden sich hinsichtlich ihrer Steuerpolitik, der Definition des zu versteuernden Einkommens, der Komplexität der Steuersysteme sowie des Verlaufs des Konjunkturzyklus. Dadurch entstehen beträchtliche Unterschiede im Hinblick auf die Höhe der Unternehmensteuerlücken.

3 Verfahren zur Schätzung der Unternehmensteuerlücke

Existierende Verfahren zur Schätzung der Unternehmensteuerlücke lassen sich grob in zwei Kategorien einteilen. Auf Makro-Daten basierende aggregierte Schätzungen (Top-down) und Mikro-basierte Schätzungen, die sich auf Informationen über einzelne Steuerpflichtige stützen (Bottom-up). Diese Kurzexpertise konzentriert sich auf Bottom-up Verfahren. Bottom-up Ansätze nutzen größtenteils Daten der Steueradministration, die aus Steuererklärungen und -bilanzen oder aus Betriebsprüfungen (Audits) stammen (Europäische Kommission, 2016). Verschiedene Länder haben bereits Programme eingeführt, um Steuerlücken anhand steueradministrativer Daten zu quantifizieren (vgl. Abschnitt 3.3.2 & 3.4.2). Aber auch Informationen aus den Handelsbilanzen von Unternehmen und aus anderen Quellen werden für die Schätzung der Unternehmensteuerlücke genutzt. Bei den Bottom-up Ansätzen wird die Steuerlücke in der Regel auf Basis einer Stichprobe von Unternehmen ermittelt und anschließend auf nationale Ebene hochgerechnet.

Im Folgenden werden verschiedene Verfahren vorgestellt und deren Anwendungsvoraussetzungen sowie Vor- und Nachteile diskutiert.

3.1 Regressionsbasierter Ansatz

3.1.1 Vorgehen

In der wirtschaftswissenschaftlichen Literatur gilt das Interesse zahlreicher Studien vor allem der Frage, wie hoch die Steuerlücke infolge einer steuermotivierten Verlagerung der Gewinne multinationaler Konzerne ist. Ein Großteil der Studien, die zu diesem Zweck auf Unternehmensdaten zurückgreifen, schätzt dabei ein multivariates Regressionsmodell der folgenden Form (vgl. Dharmapala, 2014):¹

$$(1) \log(y_{igct}) = \beta_1 \tau_{ct} + \beta_2 K_{igct} + \beta_3 L_{igct} + \gamma' X_{igct} + \delta' Z_{ct} + \mu_i + \pi_t + \varepsilon_{igct}$$

Der Index i kennzeichnet hier die Konzerngesellschaft, Index g den Konzern, Index c das Sitzland und Index t das Jahr, auf das sich die Beobachtung bezieht.² Die abhängige Variable ist der logarithmierte Gewinn vor Steuern, den die zum Konzern g gehörende Konzerngesellschaft i mit Sitz in Land c im Jahr t erzielt hat. Die wichtigste erklärende Variable im Regressionsmodell ist der Gewinnsteuersatz im Sitzland, τ_{ct} . Der zugehörige Parameter β_1 gibt an, um wieviel Prozent sich der von einer Konzerngesellschaft in ihrem Sitzland berichtete Vorsteuergewinn im Durchschnitt

¹ Als Urheber dieses Modells gelten Hines und Rice (1994). Die beiden Autoren schätzten allerdings eine leicht modifizierte Version des hier dargestellten Regressionsmodells.

² Im Unterschied zu der in Dharmapala (2014) dargestellten Regressionsgleichung enthält Gleichung (1) zusätzlich die Indizes g und c . Die Änderung im Hinblick auf die Notation dient lediglich dem Zweck, die verschiedenen Beobachtungsebenen in den Daten explizit hervorzuheben.

verändert, wenn der Steuersatz um einen Prozentpunkt ansteigt. Man spricht hierbei von der Steuer-Semielastizität des Gewinns. Gelegentlich wird anstelle des Steuersatzes im Sitzland auch die Differenz aus dem Steuersatz im Sitzland und dem Durchschnitt der Steuersätze in allen anderen Sitzländern des Konzerns g als erklärende Variable verwendet. Dies ändert jedoch grundsätzlich nichts an der Interpretation des Regressionskoeffizienten β_1 .

Durch das Hinzufügen weiterer erklärender Variablen (Kontrollvariablen) zum Regressionsmodell lässt sich die Steuer-Semielastizität unter der Bedingung ermitteln, dass andere Faktoren, die die globale Verteilung des Vorsteuergewinns eines Konzerns beeinflussen, konstant gehalten werden (*Ceteris-paribus*-Bedingung). K_{igt} misst in diesem Zusammenhang die Ausstattung einer Konzerngesellschaft mit dem Produktionsfaktor Kapital, L_{igt} steht für den Produktionsfaktor Arbeit. Als Indikator für die Ausstattung mit dem Produktionsfaktor Kapital wird häufig das (Sach-)Anlagevermögen herangezogen, für den Produktionsfaktor Arbeit die Lohnsumme oder die Anzahl an Beschäftigten. Zusammengenommen bilden diese beiden Variablen ab, wie stark das Ausmaß an realwirtschaftlicher Aktivität im Sitzland ist. X_{igt} und Z_{ct} sind Vektoren, die Variablen beinhalten, welche die Konzerngesellschaft bzw. deren Sitzland näher beschreiben und die globale Gewinnverteilung beeinflussen können. Auf Sitzlandebene werden häufig das reale Bruttoinlandsprodukt pro Kopf als Maß für das Produktivitätsniveau im Land sowie die Bevölkerungsgröße als Maß für die Bedeutung des Landes als Absatzmarkt berücksichtigt. μ_i ist ein Term, mit dem sich zeitinvariante Profitabilitätsunterschiede zwischen Konzerngesellschaften messen lassen, die beispielsweise auf unterschiedliche Geschäftsaktivitäten zurückzuführen sein können. π_t misst Veränderungen der Profitabilität im Zeitablauf, die sämtliche Konzerne und deren Gesellschaften betreffen (beispielsweise infolge technologischen Fortschritts oder konjunktureller Schwankungen).

Auf Basis des Regressionsmodells lässt sich der von Konzerngesellschaft i im Sitzland c ausgewiesene Gewinn vor Steuern in zwei Komponenten zerlegen. Die eine Komponente misst den Teil des Gewinns, der sich auf die Ausprägungen der Kontrollvariablen zurückführen lässt. Die Größe dieser Komponente wird demnach bestimmt durch das Ausmaß an realwirtschaftlicher Aktivität – gemessen durch den Kapital- und Arbeitseinsatz – sowie durch Eigenschaften der Konzerngesellschaften und des Sitzlandes, die von der Höhe des Gewinnsteuersatzes unabhängig sind (das betrifft vor allem Unterschiede im Hinblick auf das Produktivitätsniveau und die Marktgröße). Diese Komponente wird häufig als „realer“ Gewinn bezeichnet. Die zweite Komponente umfasst den „verlagerten“ Gewinn. Dieser ist unabhängig von den Ausprägungen der Kontrollvariablen und lässt sich allein durch Unterschiede in den Gewinnsteuersätzen zwischen den Sitzländern eines Konzerns erklären, weshalb er als Ergebnis von Steuergestaltungspraktiken interpretiert werden kann.

Wurde die Steuer-Semielastizität geschätzt, kann diese im nächsten Schritt genutzt werden, um den Verlust an Bemessungsgrundlage infolge steuermotivierter Gewinnverlagerung und die daraus resultierende Steuerlücke zu berechnen. Um dies zu tun, muss zunächst ermittelt werden, wie die Gewinne multinationaler Konzerne global verteilt wären, gäbe es keine steuermotivierte

Gewinnverlagerung. Diese kontrafaktische globale Gewinnverteilung wird zumeist auf Grundlage einer einfachen Simulation ermittelt, bei der die Steuersätze sämtlicher Länder der Welt auf ein einheitliches Niveau festgesetzt werden.³ Auf Basis der geschätzten Steuer-Semielastizität der Gewinne wird dabei die Veränderung der nationalen Bemessungsgrundlagen berechnet, die aus der Anpassung der Steuersätze resultiert:

$$\Delta_{ct} = \sum_i y_{igct} (\tau^* - \tau_{ct}) \hat{\beta}_1$$

Δ_{ct} misst die geschätzte Veränderung der Bemessungsgrundlage für die Gewinnbesteuerung in Land c und Jahr t . Um diese zu ermitteln, wird der von Konzerngesellschaft i in Land c und Jahr t berichtete Gewinn y_{igct} mit der Differenz aus dem Einheitssteuersatz τ^* und dem tatsächlichen Steuersatz τ_{ct} sowie der geschätzten Steuer-Semielastizität $\hat{\beta}_1$ multipliziert. Auf diese Weise erhält man für jede Konzerngesellschaft eine separate Schätzung für die Veränderung der Bemessungsgrundlage infolge einer Angleichung der nationalen Steuersätze. Summiert man diese Schätzungen für alle Konzerngesellschaft mit Sitz im selben Land auf, erhält man die geschätzten Veränderungen der nationalen Bemessungsgrundlagen. Länder, deren Gewinnsteuersatz geringer ist als der Einheitssteuersatz, erleiden infolge der Vereinheitlichung der Steuersätze einen Verlust an Bemessungsgrundlage (Δ_{ct} ist negativ). In Ländern, deren Steuersatz höher ist, nimmt die Bemessungsgrundlage zu (Δ_{ct} ist positiv).

Grundsätzlich ließe sich der Einheitssteuersatz τ^* so wählen, dass der Saldo der länderspezifischen Veränderungen der Bemessungsgrundlage null beträgt. Da ein überproportional hoher Anteil der Gewinne multinationaler Konzerne in Ländern mit eher geringem Steuersatz verbucht ist, muss der gewählte Einheitssteuersatz dafür allerdings ebenfalls gering sein. Viele Autor*innen wählen stattdessen jedoch einen Einheitssteuersatz, der höher ist und zu einem negativen Saldo führt (der geschätzte Verlust an Bemessungsgrundlage in den Niedrigsteuerländern infolge der Vereinheitlichung der Steuersätze ist größer als der Hinzugewinn an Bemessungsgrundlage in den Hochsteuerländern). In diesem Fall muss der Saldo noch unter den Ländern aufgeteilt und die geschätzten Veränderungen der nationalen Bemessungsgrundlagen entsprechend korrigiert werden. Für die Aufteilung werden zumeist realwirtschaftliche Indikatoren verwendet. Clausing (2009, 2011) beispielsweise teilt den Saldo auf Basis unternehmensinterner Transaktionsvolumina auf. Die OECD (2015) nutzt als Verteilungsschlüssel den Wert des Anlagevermögens. Fuest et al. (2022) nutzen drei unterschiedliche Verteilungsschlüssel: Den Wert des Sachanlagevermögens, die Anzahl an Beschäftigten und einen Realindikator.

³ Manche Autor*innen gleichen dabei das Steuerniveau sämtlicher Länder dem Steuerniveau desjenigen Landes an, für das die Steuerlücke berechnet werden soll (z.B. Clausing, 2009; 2011). Andere Autor*innen wählen den Einheitssteuersatz nach anderen Kriterien. Die OECD beispielsweise verwendet als erklärende Variable anstelle des Steuersatzes im Sitzland die Differenz aus dem Steuersatz im Sitzland und dem Durchschnitt der Steuersätze in allen anderen Ländern, in denen ein Konzern Gesellschaften hat (OECD, 2015; Johansson, 2017). Im Rahmen der Simulation werden diese Differenzen auf null gesetzt. Fuest et al. (2022) schätzen einen nicht-linearen Zusammenhang zwischen den Vorsteuergewinnen und Steuersätzen. Dabei stellen die Autoren fest, dass die Steuer-Semielastizität der Gewinne ab einem Steuersatzniveau von 20 Prozent gegen null tendiert. Entsprechend setzen die Autoren den Einheitssteuersatz auf 20 Prozent fest.

Im letzten Schritt werden die geschätzten (und ggf. korrigierten) Veränderungen der nationalen Bemessungsgrundlagen für die Gewinnbesteuerung mit den nationalen Gewinnsteuersätzen multipliziert, um eine Schätzung für die aus steuermotivierter Gewinnverlagerung resultierende Steuerlücke zu erhalten.

3.1.2 Forschungsergebnisse

Die Frage, wie sensitiv die Gewinne multinationaler Konzerne auf Steuersatzunterschiede zwischen Ländern reagieren, ist seit mittlerweile mehr als 30 Jahren Gegenstand empirischer Forschungsarbeiten. Beer et al. (2020a) sowie Heckemeyer und Overesch (2017) identifizieren insgesamt 37 bzw. 27 Studien, in denen die Steuer-Semielastizität der Gewinne multinationaler Konzerne auf Basis von Gleichung (1) oder einer modifizierten Version davon empirisch bestimmt wurde. Auf Grundlage von Meta-Analysen ermitteln sie eine durchschnittliche Steuer-Semielastizität der Gewinne in Höhe von -1 Prozent bzw. $-0,8$ Prozent. Dies bedeutet, dass ein Land im Durchschnitt einen Verlust an Bemessungsgrundlage für die Gewinnbesteuerung in Höhe von einem Prozent bzw. $0,8$ Prozent erleidet, wenn es den Gewinnsteuersatz um einen Prozentpunkt erhöht.

Beer et al. (2020a) und Heckemeyer und Overesch (2017) zeigen allerdings auch, dass die Höhe der geschätzten Steuer-Semielastizität der Gewinne multinationaler Konzerne stark von der Spezifikation des Regressionsmodells sowie der Definition einzelner Variablen abhängen kann. Da die Steuer-Semielastizität ein wichtiger Baustein für die Schätzung der Steuerlücke ist, dürfte ihre geschätzte Höhe ebenso sensitiv sein. In den letzten Jahren gab es darüber hinaus Kritik an der Annahme, die Steuer-Semielastizität sei über sämtliche Länder hinweg konstant. Diese Annahme wird als unzutreffend angesehen, lasse doch die besondere Bedeutung von Steueroasen darauf schließen, dass die Steuer-Semielastizität betragsmäßig größer wird, je geringer der Steuersatz in einem Land ist (Dowd et al., 2017). Mit anderen Worten: Hebt ein Niedrigsteuerland seinen Steuersatz an, so hat das einen (relativ) größeren Gewinnabfluss zur Folge, als wenn ein Hochsteuerland seinen Steuersatz anhebt. In neueren empirischen Studien wurde diesem Umstand auf verschiedene Weise Rechnung getragen:

- Dowd et al. (2017), die die Steuer-Semielastizität der Gewinne US-amerikanischer multinationaler Konzerne schätzen, fügen Gleichung (1) einen quadratischen Steuerterm hinzu. In einer weiteren Spezifikation interagieren sie ihre Steuersatzvariable mit einer binären Variable, die den Wert 1 annimmt, wenn ein Land einen besonders niedrigen Steuersatz hat (Steuersatz kleiner als das 5%-Perzentil).
- Bratta et al. (2021) schätzen eine quadratische und eine kubische Spezifikation auf Basis von Daten aus *Country-by-Country*-Berichten von multinationalen Großkonzernen mit einem Steuersitz in Italien.
- Garcia-Bernado und Jansky (2020) logarithmieren die Steuersatzvariable.
- Fuest et al. (2022) schätzen eine sog. *Restricted-Cubic-Spline*-Spezifikation, bei der die Länder in Abhängigkeit von der Höhe des Steuersatzes in fünf Gruppen eingeteilt werden.

Für jede Gruppe wird dann eine kubische Spezifikation geschätzt. Datengrundlage sind *Country-by-Country*-Berichte von in Deutschland aktiven multinationalen Großkonzernen.

Alle vier Studien kommen dabei zu dem Ergebnis, dass die Steuer-Semielastizität betragsmäßig größer ist, je geringer das Steuerniveau in einem Land ist. Wie groß genau hängt allerdings stark von der gewählten funktionalen Form ab. Dowd et al. (2017) ermitteln auf Basis einer quadratischen Spezifikation eine minimale Steuer-Semielastizität von -4 , Fuest et al. (2022) auf Grundlage einer *Restricted-Cubic-Spline*-Spezifikation von -13 .

Von den 37 in Beer et al. (2020a) zitierten Studien schätzt allerdings nur ein geringer Teil auch die Steuerlücke. Ein Grund dafür könnte sein, dass die in vielen Studien verwendeten Daten die Aktivitäten multinationaler Konzerne nur unvollständig abbilden, was die Schätzung der Steuerlücke erschwert (siehe Abschnitt 3.1.3). 20 der 37 in Beer et al. (2020a) zitierten Studien nutzen für ihre Schätzung öffentlich zugängliche Bilanzinformationen, die von *Bureau van Dijk* gesammelt und aufbereitet werden. Diese Daten erfassen jedoch nicht die Aktivitäten sämtlicher Gesellschaften multinationaler Konzerne. Darüber hinaus beinhalten die Daten überhaupt keine Informationen zu den Aktivitäten in vielen Niedrigsteuerländern (Tørsløv et al., 2018).

Clausing (2009, 2011) verwendet den oben beschriebenen Ansatz, um die aus steuermotivierter Gewinnverlagerung resultierende Steuerlücke für die USA zu berechnen. Für ihre Analyse nutzt die Autorin Daten des *Bureau of Economic Analysis*, welche Informationen zu den globalen Aktivitäten multinationaler US-Konzerne umfassen. Die Ergebnisse der beiden Studien legen nahe, dass die USA durch steuermotivierte Gewinnverlagerung einen Verlust an Einnahmen aus Unternehmensteuern von bis zu 35 Prozent erleiden.

Fuest et al. (2022) verwenden Informationen aus den *Country-by-Country*-Berichten sämtlicher in Deutschland aktiver multinationalen Konzerne, um die Aufkommensverluste infolge steuermotivierter Gewinnverlagerung für Deutschland zu berechnen. Die Autoren kommen zu dem Ergebnis, dass Deutschland pro Jahr zwischen 1,6 und 2,6 Milliarden Euro an Aufkommen durch steuermotivierte Gewinnverlagerung einbüßt. Dabei ist allerdings anzumerken, dass der in Fuest et al. (2022) verwendete Datensatz ausschließlich Informationen zu den in Deutschland aktiven multinationalen Großkonzernen (globaler Umsatz beträgt mindestens 750 Millionen Euro pro Jahr) beinhaltet. Der Aufkommensverlust, der sich auf eine steuermotivierte Verlagerung der Gewinne kleinerer multinationaler Konzerne zurückführen lässt, bleibt hier unberücksichtigt.

Studien der OECD (2015) und von Johansson et al. (2017) enthalten Schätzungen darüber, wie hoch die Steuerlücke infolge steuermotivierter Gewinnverlagerung weltweit ist. Den Ergebnissen zufolge liegt der globale Verlust an Steueraufkommen zwischen 4 Prozent und 10 Prozent des weltweiten Aufkommens aus Unternehmensteuern.

3.1.3 Bewertung

Obwohl der regressionsbasierte Ansatz weit verbreitet ist, sollten dessen Ergebnisse mit Vorsicht interpretiert werden. Ein wichtiger Grund hierfür ist die oben angesprochene Sensitivität der Ergebnisse im Hinblick auf die Spezifikation des Regressionsmodells und die Definition der darin verwendeten Variablen. Die Ergebnisse der Schätzung von Steuer-Semielastizität und Steuerlücke sind dadurch ein Stück weit immer auch getrieben von den verwendeten Daten sowie den Entscheidungen der Studienautor*innen im Hinblick auf das verwendete empirische Modell.

Bei der Schätzung des durch Gleichung (1) repräsentierten Regressionsmodells wird ferner angenommen, dass es jenseits des Steuersatzes und der Kontrollvariablen keine anderen Faktoren gibt, die die globale Verteilung der Konzerngewinne systematisch beeinflussen und gleichzeitig mit dem Steuersatz oder den Kontrollvariablen korrelieren. Ist diese Annahme verletzt, kann das dazu führen, dass die geschätzte Steuer-Semielastizität und damit auch die Steuerlücke verzerrt sind. Riedel (2018) nennt einige Beispiele für solche Verletzungen.

Eine wichtige Voraussetzung, um die Steuerlücke auf Grundlage von Unternehmensdaten schätzen zu können, ist, dass der verwendete Datensatz die weltweiten Aktivitäten multinationaler Konzerne vollständig erfasst. Insbesondere muss der Datensatz Informationen über sämtliche (in- und ausländische) in einem Land aktiven multinationalen Konzerne beinhalten. Darüber hinaus muss der Datensatz sowohl die Inlands- als auch die Auslandsaktivitäten der Konzerne vollständig abbilden, und das für jedes Sitzland separat. Andernfalls besteht die Gefahr, dass die geschätzte Steuer-Semielastizität der Vorsteuergewinne und die Steuerlücke eine Verzerrung aufweisen. Umfasst der Datensatz nicht alle in einem Land aktiven multinationalen Konzerne, müssen die Ergebnisse, die auf Basis der Stichprobe erzielt werden, zudem hochgerechnet werden. Eine solche Hochrechnung führt im Allgemeinen jedoch nur dann zu validen Ergebnissen, wenn angenommen werden kann, dass die Konzerne, für die sämtliche relevanten Informationen vorhanden sind, eine repräsentative Stichprobe aller im Land aktiven Konzerne darstellen.

Abschließend sei noch einmal darauf hingewiesen, dass die hier skizzierte Methode lediglich dem Zweck dient, jenen Teil der Steuerlücke zu identifizieren, der sich auf steuermotivierte Gewinnverlagerung durch multinationale Konzerne zurückführen lässt. Über die Höhe der Steuerlücke im engeren Sinn gibt dieses Verfahren keinen Aufschluss.

Ferner sei angemerkt, dass die Existenz einer aus steuermotivierter Gewinnverlagerung resultierenden Steuerlücke kein Indiz für illegales Verhalten ist. Vielmehr ist davon auszugehen, dass multinationale Konzerne legale Wege und Schlupflöcher nutzen, um Gewinne in Niedrigsteuermölandern zu verlagern. Diese Schlupflöcher können durch Lücken in nationalen Steuergesetzen entstehen oder durch Inkonsistenzen in Doppelbesteuerungsabkommen. Möglich ist auch, dass multinationale Konzerne ihre Aktivitäten global so organisieren, dass sie nationale Steuerprivilegien für bestimmte Einkunftsarten optimal ausnutzen.

3.2 Matching-Ansatz

3.2.1 Vorgehen

Der regressionsbasierte Ansatz für die Berechnung der Unternehmensteuerlücke ist dadurch gekennzeichnet, dass die Profitabilität der Gesellschaften multinationaler Konzerne über die Sitzländer hinweg verglichen wird. Die Steuer-Semielastizität, die für die Berechnung der Steuerlücke entscheidend ist, wird bestimmt, indem man die globale Verteilung der Gewinne über Hoch- und Niedrigsteuerländer mit der globalen Verteilung realwirtschaftlicher Aktivität vergleicht.

Ein anderer Ansatz für die Messung der Steuerlücke infolge steuermotivierter Gewinnverlagerung ist, multinationale Konzerne mit ausschließlich national agierenden Unternehmen zu vergleichen. Dahinter steht die Annahme, dass Unternehmen, die ausschließlich national tätig sind, keine Möglichkeit haben, Gewinne ins Ausland zu verlagern. Beim Vergleich konzentriert man sich dabei auf die Steuerzahlungen der beiden Unternehmensgruppen. Um sicherzustellen, dass ein Vergleich auch zulässig ist bzw. dass Unterschiede im Hinblick auf die Steuerzahlungen, die die beiden Unternehmenstypen leisten, nicht durch andere Faktoren bedingt sind, wird ein sogenanntes Matching-Verfahren angewendet. Beim Matching wird für jede Gesellschaft eines multinationalen Konzerns, die in dem Land, für das die Steuerlücke berechnet werden soll, ansässig ist, ein (oder ggf. mehrere) nationales Unternehmen ausgewählt, das der Konzerngesellschaft im Hinblick auf Eigenschaften, die einen Einfluss auf die Höhe der Steuerzahlungen haben, möglichst ähnlich ist. Man spricht hierbei auch von „statistischen Zwillingen“. Für die Auswahl der statistischen Zwillinge gibt es verschiedene Methoden. Lange Zeit war in der wirtschaftswissenschaftlichen Literatur die sogenannte *Propensity-Score*-Methode weit verbreitet. Hierbei werden statistische Zwillinge auf Basis eines binären Regressionsmodells geschätzt. Die binäre abhängige Variable zeigt an, ob es sich bei dem betrachteten Unternehmen um die Gesellschaft eines multinationalen Konzerns handelt oder nicht. Unternehmenscharakteristika, die die Höhe der Steuerzahlungen beeinflussen (können), werden dem Regressionsmodell als erklärende Variablen hinzugefügt. Alternativ lassen sich statistische Zwillinge auf Basis von Distanzmaßen (z.B. Euklidische Distanz, Mahalanobis-Distanz) identifizieren. Aktuell gewinnen jedoch Matching-Verfahren an Bedeutung, die auf Methoden maschinellen Lernens aufbauen. Diese weisen gegenüber dem *Propensity-Score*-Matching und dem Matching auf Basis eines Distanzmaßes mehrere Vorteile auf.

Wurde für jede im Inland ansässige Gesellschaft eines multinationalen Konzerns (mindestens) ein statistischer Zwilling aus dem Pool nationaler Unternehmen identifiziert, lässt sich die Steuerlücke berechnen, indem man für jedes Zwillingspaar die Differenz zwischen der Steuerzahlung der Konzerngesellschaft und der Steuerzahlung des nationalen Unternehmens berechnet. Anschließend summiert man die Differenzen über alle Zwillingspaare hinweg auf. Die Summe entspricht der geschätzten Steuerlücke.

3.2.2 Forschungsergebnisse

Aktuell existieren nach unserem Wissen lediglich drei Studien, die die Unternehmensteuerlücke auf Basis eines Matching-Verfahrens ermitteln. Alle drei Studien schätzen dabei lediglich die Steuerlücke für eine durchschnittliche inländische Gesellschaft eines multinationalen Konzerns, nicht jedoch die Höhe der Steuerlücke auf nationaler Ebene.

Egger et al. (2010) nutzen Daten über europäische Firmen bzw. Gesellschaften multinationaler Konzerne mit Sitz in Europa für ihren Vergleich. Die Autoren berichten, dass nationale Unternehmen in europäischen Hochsteuerländern im Schnitt mehr als 30 Prozent mehr an Steuern zahlen als vergleichbare inländische Gesellschaften multinationaler Konzerne. Für Deutschland berichtet Finke (2013) eine Steuerlücke von 27 Prozent in 2007. Diese verringerte sich allerdings in den Folgejahren, was die Autorin auf die Steuerreform im Jahr 2008 zurückführt. Hansson et al. (2019) konzentrieren sich in ihrer Analyse auf Unternehmen aus dem verarbeitenden Gewerbe in Schweden. Die Autor*innen ermitteln eine Steuerlücke in Höhe von 19 Prozent.

3.2.3 Bewertung

Die größte Herausforderung bei der Anwendung des Matching-Ansatzes für die Schätzung der Unternehmensteuerlücke besteht darin, sämtliche Eigenschaften zu identifizieren und zu operationalisieren, die einen Einfluss auf die Höhe der Steuerschuld haben können. Ist die Liste an Eigenschaften, die für die Identifikation statistischer Zwillinge genutzt wird, unvollständig, besteht die Gefahr, dass es bei der Schätzung der Steuerlücke zu Verzerrungen kommt.

Im Hinblick auf die Anforderungen an die Datengrundlage hat der Matching-Ansatz gegenüber dem regressionsbasierten Verfahren einen wichtigen Vorteil: Es werden lediglich Informationen zu den Inlandsaktivitäten multinationaler Konzerne benötigt, nicht jedoch über deren Auslandsaktivitäten.

Mit dem regressionsbasierten Verfahren gemein hat der Matching-Ansatz, dass auch hier der Fokus auf dem Beitrag steuermotivierter Gewinnverlagerung zur Steuerlücke liegt. Der Beitrag anderer Faktoren bleibt unberücksichtigt. Somit gibt auch dieser Ansatz keinen Aufschluss über die Höhe der Unternehmensteuerlücke im engeren Sinn.

3.3 Randomisierte Betriebsprüfungen

3.3.1 Vorgehen

Bei der randomisierten Betriebsprüfung werden die Steuererklärungen und Geschäftskonten einer zufällig ausgewählten Stichprobe (Umfang n) an Unternehmen überprüft und auf Unregelmäßigkeiten, welche auf Steuerhinterziehung deuten, untersucht. Bei Abschluss der Prüfung muss das Unternehmen (gekennzeichnet durch den Index k), wenn das zu versteuernde Einkommen zu gering angegeben wurde, eine Steuernachzahlung leisten oder erhält eine Rückzahlung bei zu hoch angesetztem zu versteuerndem Einkommen. Die Summe der Differenzen zwischen der ursprünglichen Steuerzahlung (s_k^A) und der Steuerzahlung nach Durchführung der Prüfung

(s_k^B)) entspricht der Stichproben-Steuerlücke. Durch Hochrechnung erhält man eine Schätzung für die Höhe der nationalen Unternehmensteuerlücke. Hierfür wird angenommen, dass die Höhe der Stichproben-Steuerlücke proportional ist zum Anteil der geprüften Unternehmen an allen Unternehmen (deren Anzahl beträgt N):

$$U = \frac{N}{n} * \sum_{k=1}^n s_k^B - s_k^A$$

Für eine fundierte Hochrechnung muss die gewährleistet sein, dass die gezogene Stichprobe repräsentativ ist für die Grundgesamtheit der Unternehmen. Anderenfalls können die Ergebnisse verzerrt sein. Die Repräsentativität der Stichprobe lässt sich jedoch nicht immer erkennen (Gemmel & Hasseldine, 2012). Manche Steuerbehörden, wie die kanadische *Canadian Revenue Agency* (CRA) und die britische *Her Majesty's Revenue and Customs* (HMCR), nutzen eine stratifizierte Randomisierung, die gewährleistet, dass Unternehmen aus allen wichtigen Gruppen (z.B. Industrien, gekennzeichnet durch den Index i) vertreten sind. Für den Fall, dass Teilgruppen in der Stichprobe überrepräsentiert sind, um aussagekräftige Steuerlückenergebnisse für jede Teilgruppe zu erhalten, wird die Hochrechnung mit Gewichten (w_i) an die Grundgesamtheit angepasst.

$$U = \frac{N}{n} * \sum_{k=1}^n w_i (s_{k,i}^B - s_{k,i}^A)$$

Weiterhin muss gewährleistet sein, dass im Rahmen der Steuerprüfung sämtliche Fälle, in denen es eine Abweichung gibt zwischen der geschuldeten Steuer und der gezahlten Steuer, auch aufgedeckt werden und die Höhe der Abweichung korrekt ermittelt wird. Falls dies nicht der Fall ist, beispielsweise weil nur ein Teil der Geschäftsaktivitäten geprüft wird oder nicht alle Steuerprüfer*innen Abweichungen korrekt identifizieren, wird die Steuerlücke unterschätzt. Um dem entgegenzusteuern, können Nichterkennungs-Multiplikatoren und Korrekturverfahren eingesetzt werden. Mögliche Methoden sind: Drittparteien-Vergleiche, die *Detection-Controlled-Estimation*-Methode (DCE) und Zweitbewertungen.⁴

Betriebsprüfungen können sich aufgrund von möglichen Einsprüchen oder gerichtlichen Anfechtungen teilweise über Jahre hinziehen (Bloomquist et al., 2014; HMCR, 2021). Da es das Ziel der Schätzung der Steuerlücke ist, einen Richtwert für das Steuereinnahmepotenzial zu erhalten, sollten die Prüfungen zum Zeitpunkt der Schätzung vollständig abgeschlossen sein. Schätzungen der Steuerlücke, wie die des US-amerikanischen *Internal Revenue Service* (IRS), werden aus

⁴ Die DCE wird vom US-amerikanischen *Internal Revenue Service* (IRS) zum Berechnen der Einkommenssteuerlücke verwendet. Hierbei wird angenommen, dass sich Prüfer*innen hinsichtlich ihrer Effizienz – entdeckte Einkommenssteuerabweichungen in Relation zu allen Abweichungen in der Steuerklärung – unterscheiden. Um für diesen Faktor zu korrigieren, wird das Ergebnis der Betriebsprüfung auf das Ergebnis, welches der/die bestqualifizierte Prüfer*in erzielt hätte, hochskaliert (Erard & Feinstein, 2011). Der Drittparteien-Vergleich wird von der britischen HMCU genutzt. Hierbei wird die Betriebsprüfung zuerst ohne die Einkommensinformationen der Drittpartei durchgeführt. Nachträglich wird dieses Ergebnis verglichen mit dem Betriebsprüfungsergebnis, welches entsteht, wenn zusätzlich Informationen der Drittpartei über sämtliche Einkommensquellen hinzugezogen werden. Der Unterschied zwischen den Ergebnissen der beiden Betriebsprüfungen entspricht dem Nichterkennungsanteil.

diesem Grund erst Jahre nach dem betrachteten Steuerjahr veröffentlicht. Um die Ergebnisse frühzeitiger veröffentlichen zu können, nutzt die britische HMCR die Erfolgsrate ähnlicher offener Fälle aus den vorangegangenen Jahren, um die Steuererträge zu schätzen.

3.3.2 Ergebnisse

Randomisierte Steuerprüfungen gelten als Goldstandard, um die Steuerlücke zu schätzen und Einblicke in die Ursachen für die Existenz der Steuerlücke zu erhalten.⁵ Seit einigen Jahren steigt deshalb die Anzahl an durchgeführten Betriebsprüfungen sowie der wissenschaftlichen Studien, welche die Steuerlücke auf Basis von Informationen aus Betriebsprüfungen schätzen und die Vorteile dieses Verfahrens unterstreichen (Gemmel & Hasseldine, 2012; Feinstein, 1999; Europäische Kommission, 2018). Laut einer Umfrage der OECD (2017) gaben im Jahr 2015 insgesamt 35 Länder an, randomisierte Betriebsprüfungen durchzuführen. 12 dieser Länder nutzen die daraus gewonnenen Informationen, um Steuerlücken zu schätzen.⁶ Die OECD fördert zusätzlich seit 2019 den Austausch von Steuerlückenschätzungen innerhalb der Länder im Rahmen des internationalen Forums *Advanced Tax Gap Analysis Community of Practice*. In Europa wird die Steuerlücke am häufigsten für Umsatzsteuern geschätzt (Europäische Kommission, 2016). Laut einer Umfrage der Europäischen Kommission hat jedoch 2017 ein Großteil der Länder bereits Unternehmensteuerprüfungen durchgeführt.⁷ Sieben der befragten Länder (Belgien, Bulgarien, Dänemark, Italien, die Niederlande, Schweden, Finnland) gaben an, entweder risikobasierte oder randomisierte Betriebsprüfungen zur Berechnung der Steuerlücke durchzuführen oder dies zu planen (Europäische Kommission, 2018). Außerhalb Europas ist das *National Research Program* (NRP) der US-amerikanischen Steuerbehörde IRS bekannt für seine regelmäßigen Steuerlückenschätzungen auf Basis randomisierter und risikobasierter Steuerprüfungen. Im Vergleich zur Einkommenssteuerlücke basiert die Unternehmensteuerlückenschätzung des IRS jedoch auf risikobasierten Betriebsprüfungen.

Dänemark: Für die Schätzung der Unternehmensteuerlücke führt die dänische Steuerbehörde (SKAT) sowohl randomisierte (alle zwei Jahre seit 2006) als auch risikobasierte Betriebsprüfungen durch. Im Jahr 2012 enthielt die randomisierte Stichprobe 1.000 Unternehmen. Da in Dänemark zu diesem Zeitpunkt insgesamt 220.000 Unternehmen gemeldet waren, wird für die Hochrechnung der Ergebnisse auf nationale Ebene jedem Unternehmen ein Gewicht von 220 zugeteilt. In diesem Jahr ergab die Hochrechnung eine geschätzte Unternehmensteuerlücke von 4,4 Milliarden dänischen Kronen, was 20 Prozent der tatsächlichen Steuerschuld entsprach. 56 Prozent der Steuererklärungen waren fehlerhaft (im Jahr 2006 waren es 42 Prozent, im Jahr 2008 52 Prozent). Die Hälfte der Steuerlücke ließ sich auf bewusste Steuerhinterziehung von insgesamt 10 Prozent der Unternehmen zurückführen. Die Ergebnisse aus verschiedenen Jahren zeigen, dass

⁵ Siehe IRS S.4.: <https://home.treasury.gov/system/files/136/The-American-Families-Plan-Tax-Compliance-Agenda.pdf#page=5>

⁶ https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.oecd-ilibrary.org%2Fdocservers%2Ftax_admin-2017-table198-en.xlsx%3Fexpires%3D1652423988%26id%3Did%26ac-cname%3Docid56012597%26checksum%3D578201FA9D47FE801C43AAAA101686E3&wdOrigin=BROWSELINK

⁷ Siehe Tabelle 5.1 in Europäische Kommission (2018).

Unternehmen, welche zur Hotel- oder Restaurant-Branche gehören, am häufigsten Steuern vermeiden. Unternehmen der Handels- oder Landwirtschaftsindustrie haben hingegen den größten Anteil an der Steuerlücke. Die Steuerprüfungen dauern im Schnitt ca. 18 Monate. Die Auswertung der Daten und die Hochrechnung nimmt ein weiteres Jahr in Anspruch (Søndergaard, 2018; Europäische Kommission, 2018).

Kanada: Die kanadische Bundesbehörde CRA gründete im April 2016 eine Einheit, deren Aufgabe die Schätzung von Steuerlücken für verschiedene Steuerarten ist. Im Zeitraum von 2016 bis 2019 wurden umfangreiche Studien veröffentlicht, darunter auch eine zur Unternehmensteuerlücke im Jahr 2014 (CRA, 2019). Für die Schätzung wurden randomisierte Betriebsprüfungen bei mittelgroßen und Kleinunternehmen (KMU)⁸ sowie – aufgrund ihrer geringen Anzahl⁹ – risikobasierte Prüfungen bei Großunternehmen durchgeführt. Für KMU nutzt die CRA eine stratifizierte Randomisierung, um alle Wirtschaftszweige ausreichend abzudecken. Die geschätzte Steuerlücke für KMU beträgt zwischen 1,6-3,5 Milliarden kanadische Dollar, was ca. 4 Prozent bis 6 Prozent der gesamten Unternehmensteuereinnahmen ausmacht. Die CRA nutzt keine Nichtentdeckungs-Multiplikatoren, weshalb die Autor*innen der Studie einräumen, dass die tatsächliche Steuerlücke vermutlich größer ist (CRA, 2019).

Großbritannien: In Großbritannien führt die nicht-ministerielle Abteilung HMRC Betriebsprüfungen durch, welche für die Schätzung der jährlichen Steuerlücke genutzt werden.¹⁰ Hierzu gehört auch die Unternehmensteuerlücke für KMU und Großunternehmen (Bloomquist et al., 2014). Bei Kleinunternehmen werden randomisierte Betriebsprüfungen durchgeführt, für mittelgroße und Großunternehmen risikobasierte Betriebsprüfungen. Die Randomisierung bei den Kleinunternehmen ist stratifiziert hinsichtlich der jährlichen Handelsumsätze der Unternehmen. Die prozentuale Unternehmensteuerlücke für das Steuerjahr 2019-2020 beträgt laut Schätzung insgesamt 7,2 Prozent.¹¹ Bei Großunternehmen beträgt der Fehlbetrag 2,6 Prozent (eine Milliarde britische Pfund) der tatsächlichen Steuerschuld. Mittlere und Kleinunternehmen weisen eine Unternehmensteuerlücke von 8,9 Prozent (1,4 Milliarden Pfund) bzw. 14,7 Prozent (3,4 Milliarden Pfund) auf. Im Steuerjahr 2005-2006 lag die geschätzte Steuerlücke für die drei Unternehmensgruppen noch bei 8,7 Prozent (Großunternehmen) bzw. 13,3 Prozent (mittelgroße Unternehmen) bzw. 18,9 Prozent (Kleinunternehmen). Um nicht aufgedeckte Fehlbeträge in der Hochrechnung abzubilden, nutzt die HMRC einen Nichtentdeckungs-Multiplikator. Zudem veröffentlicht die HMCR ihre Ergebnisse, noch bevor alle offenen Untersuchungen abgeschlossen sind. Da die Un-

⁸ KMU sind in dieser Klassifikation Unternehmen mit Bruttoerträgen von höchstens 20 (in manchen Industrien 50) Millionen kanadische Dollar im Jahr.

⁹ 1% aller Unternehmen fallen in die Gruppierung Großunternehmen. Sie machen jedoch 54% der Unternehmensteuereinnahmen aus.

¹⁰ <https://www.gov.uk/government/statistics/measuring-tax-gaps>

¹¹ Die ursprüngliche Berechnung von 7,6 Prozent (HMRC, 2021) wurde nachträglich korrigiert: <https://www.gov.uk/government/statistics/measuring-tax-gaps/measuring-tax-gaps-2021-edition-tax-gap-estimates-for-2019-to-2020#corporation-tax-1>

ternehmensteuerlücke des HMCR jedoch das Nettoergebnis widerspiegelt – die auf alle Unternehmen hochgerechnete Steuerlücke abzüglich der von den geprüften Unternehmen erhaltenen Nachzahlungen – werden für die offenen Untersuchungen die erwarteten Steuererträge auf Basis der Erfolgsrate vorheriger ähnlicher Untersuchungen geschätzt. Dies führt zu Korrekturen der Ergebnisse in den folgenden Jahren (HMCR, 2021; HMCR, 2022).

Australien: Die australische Steuerbehörde ATO (*Australian Taxation Office*) hat für das Steuerjahr 2018/2019 bereits zum vierten Mal in Folge ihre Ergebnisse für die Schätzung der Einkommens-, Unternehmens- und Umsatzsteuerlücke sowie für die Steuerlücke weiterer Steuerarten veröffentlicht. Die Unternehmensteuerlückenschätzung findet separat für Klein-, Mittel- und Großunternehmen statt.¹² Die Schätzungen für Kleinunternehmen basieren auf randomisierten Betriebsprüfungen, für Mittel- und Großunternehmen auf risikobasierten Betriebsprüfungen. Um Schätzfehler durch Nichtentdeckung zu adressieren, wird das Ergebnis hochskaliert, basierend auf dem Mittelwert internationaler Skalierungswerte. Das neueste Steuerlückenergebnis ergibt eine Netto-Steuerlücke von 4,3 Prozent (2,6 Milliarden australische Dollar) bei Großunternehmen, 6,2 Prozent (0,8 Milliarden Dollar) für mittelgroße Unternehmen und 12,7 Prozent (12,5 Milliarden Dollar) für Kleinunternehmen.¹³

In vielen Ländern arbeitet die Steuerverwaltung mit Wissenschaftler*innen zusammen oder stellt der Wissenschaft die Informationen aus Betriebsprüfungen zur Verfügung, um zusätzliche Erkenntnisse zu gewinnen, die über die Schätzung der Höhe der Steuerlücke hinausgehen.¹⁴

Hoopes et al. (2012) analysieren anhand von Daten des IRS, ob US-amerikanische börsennotierte Unternehmen ihre Steuerplanung anpassen, wenn die Wahrscheinlichkeit, geprüft zu werden, steigt. Die Ergebnisse der Autoren legen nahe, dass eine Erhöhung der Wahrscheinlichkeit, geprüft zu werden, von 19 Prozent auf 37 Prozent den effektiven Steuersatz – also das Verhältnis von Steuerzahlungen zum Gewinn vor Steuern – um 7 Prozent erhöht (2 Prozentpunkte). Die Ergebnisse legen damit nahe, dass häufige Steuerprüfungen die Steuerlücke verringern und zu höheren Steuereinnahmen führen, sodass die Kosten von Steuerprüfungen (zum Teil) wieder eingespielt werden. Zu einem weniger positiven Ergebnis kommen DeBacker et al. (2015), die ebenfalls anhand von IRS-Daten untersuchen, wie Unternehmen ihre Steuererklärungen bis zu

¹² Die Grenzwerte des jährlichen Umsatzes zur Klassifizierung der Unternehmen sind 10 und 250 Millionen australische Dollar.

¹³ Die Informationen hierzu stammen von der Website des ATO: <https://www.ato.gov.au/About-ATO/Research-and-statistics/In-detail/Tax-gap/Australian-tax-gaps-overview/>

¹⁴ Ein Großteil der wissenschaftlichen Arbeiten, die auf Informationen aus Steuerprüfungen aufbauen, konzentrieren sich auf die Einkommensteuer. Dieser Literaturzweig fokussiert sich vor allem auf Verhaltensänderungen von Steuerpflichtigen nach erfolgter Steuerprüfung oder aber nach der Bereitstellung von Informationen über zukünftige Prüfungen im Rahmen von sogenannten randomisiert-kontrollierten Studien (Slemrod, 2019). Die Studien kommen zu dem Ergebnis, dass Steuerprüfungen und Steuerprüfungswarnungen (z.B. per Brief) das angegebene zu versteuernde Einkommen bei der folgenden Steuerklärung positiv beeinflussen (DeBacker et al. 2018; Advani et al., 2019; Bérigolo et al., 2018a; Kleven et al., 2011; Slemrod et al., 2001) und dass die Effizienz von Steuerprüfungen sich ebenfalls positiv auf das zu versteuernde Einkommen im folgenden Jahr auswirkt (Gemmel & Ratto, 2012; Beer et al., 2020b; Kasper & Alm, 2022).

zehn Jahre nach einer Betriebsprüfung anpassen. Den Ergebnissen der Studie zufolge weisen Unternehmen im Jahr nach einer Betriebsprüfung einen niedrigeren effektiven Steuersatz auf (-0,7 Prozent). Der negative Trend bleibt weitere Jahre bestehen, kehrt sich jedoch nach ein paar Jahren wieder um. Das Ergebnis lässt sich so interpretieren, dass Unternehmen nach einer Betriebsprüfung eine aggressivere Steuerplanung betreiben. Daraus schließen die Autor*innen, dass kürzlich geprüfte Unternehmen davon ausgehen, dass das Risiko, erneut geprüft zu werden, gering ist. Andererseits zahlen Unternehmen, die demselben Wirtschaftszweig angehören wie geprüfte Unternehmen, selbst jedoch nicht geprüft worden sind, im Jahr nach einer Prüfung effektiv mehr Steuern. Dies bestärkt den Verdacht, dass eine hohe Zahl an Betriebsprüfungen dazu beitragen kann, die Unternehmensteuerlücke zu schließen. Es sei jedoch darauf hingewiesen, dass die von DeBacker et al. (2015) verwendete Stichprobe nicht repräsentativ ist, da nur Unternehmen berücksichtigt wurden, die nicht mehrfach geprüft wurden.

D'Agosto et al. (2017) analysieren in Ihrer Studie den Einfluss von Betriebsprüfungen auf das Verhalten von Kleinunternehmen anhand von Informationen aus Betriebsprüfungen in Italien. Kleinunternehmen machen 98 Prozent der italienischen Unternehmen aus. Den Ergebnissen zufolge führt eine Betriebsprüfung im Durchschnitt zu einer 1,8 Prozent höheren Steuerzahlung im Prüfungsjahr und zu einem Anstieg von 3,1 Prozent im Folgejahr.

Bérgolo et al. (2018b) nutzen eine Umfrage unter Unternehmen in Uruguay, um zu erfahren, wie hoch die Wahrscheinlichkeit eingeschätzt wird, einer Betriebsprüfung unterzogen zu werden. Ihr Ergebnis zeigt, dass Unternehmen diese Wahrscheinlichkeit durchschnittlich höher einschätzen (39,5 Prozent) als sie tatsächlich ist (ca. 8 Prozent). Gleichzeitig schätzen die Unternehmen auch die Höhe der möglichen Strafzahlungen falsch ein. Die Autor*innen schließen daraus, dass es nützlich für die Steuerbehörden ist, intransparent im Hinblick auf die Wahrscheinlichkeit einer Betriebsprüfung zu sein.

3.3.3 Bewertung

Während der regressionsbasierte Ansatz und der Matching-Ansatz die Steuerlücke mit dem Aufkommensverlust infolge steuermotivierter Gewinnverlagerung durch multinationale Unternehmen gleichsetzen, ist der Fokus bei jenen Ansätzen, die die Steuerlücke auf Basis von Informationen aus Betriebsprüfungen schätzen, ein anderer. Die Steuerlücke ist hier definiert als Verlust an Aufkommen, der durch fehlerhafte und/oder falsche Angaben in den Steuererklärungen zustande kommt oder durch Zahlungsverzögerungen auf Seiten der Steuerpflichtigen. Geschätzt wird hier also das Ausmaß der Unternehmensteuerlücke gemäß der engen Definition.

Informationen aus Betriebsprüfungen für die Schätzung der Unternehmensteuerlücke heranzuziehen, bringt einige Vorteile mit sich. Erstens lassen sich durch Betriebsprüfungen die Ursachen für das Auftreten einer Steuerlücke detailliert nachvollziehen. Dadurch kann der Beitrag der einzelnen Komponenten zur Steuerlücke quantifiziert werden. Zweitens können durch eine systematische Auswertung der Informationen aus den Betriebsprüfungen typische Formen der Steuerrückbildung erkannt sowie häufige Ursachen für unbeabsichtigte Fehler in den

Steuererklärungen identifiziert werden.¹⁵ Diese Erkenntnisse können genutzt werden, um Lücken und Verbesserungspotenziale im Steuersystem und beim Steuervollzug offenzulegen, wodurch wiederum die Steuerlücke minimiert werden kann. Drittens kann die Offenlegung von (aggregierten) Ergebnissen von Betriebsprüfungen als ein wichtiges Signal für die Effektivität der Steuerverwaltung dienen und auf diesem Wege dazu beitragen, die Steuerehrlichkeit zu erhöhen.

Werden Betriebsprüfungen randomisiert durchgeführt, hat das den Vorteil, dass die Gefahr, dass es bei der Schätzung der Steuerlücke zu Verzerrungen kommt, minimiert wird. Die Randomisierung gewährleistet, dass die gezogene Stichprobe an Unternehmen repräsentativ ist für die Grundgesamtheit sämtlicher Unternehmen. Verzerrungen können lediglich dann auftreten, wenn Steuerhinterziehung nicht entdeckt bzw. fehlende/falsche Angaben durch die Prüfer*innen nicht erkannt werden. Allerdings besteht wie oben erwähnt die Möglichkeit, die Schätzungen entsprechend anzupassen und so Verzerrungen zu verringern.

Der größte Nachteil von randomisierten Betriebsprüfungen ist der damit verbundene hohe Kosten- und Zeitaufwand. Wenn die Daten einmal verfügbar sind, ist der Aufwand für die Schätzung der Steuerlücke allerdings sehr gering (Europäische Kommission, 2018).

3.4 Risikobasierte Betriebsprüfungen

3.4.1 Vorgehen

Risikobasierte Betriebsprüfungen konzentrieren sich auf Unternehmen, bei denen davon ausgegangen wird, dass Unregelmäßigkeiten in der Steuererklärung mit höherer Wahrscheinlichkeit auftreten. Dabei werden Kriterien identifiziert, welche auf ein erhöhtes Risiko hinweisen. Aufgrund des Fokus auf Unternehmen, bei der die Wahrscheinlichkeit einer Nachzahlung als hoch eingestuft wird, ist die risikobasierte Stichprobe nicht repräsentativ. Würde man das durchschnittliche Mehrergebnis dieser Unternehmen auf nationale Ebene hochrechnen, würde man die Unternehmensteuerlücke vermutlich überschätzen. Da es jedoch keine Vergleichsdaten von anderen Unternehmen gibt, müssen die zu erwartenden Verzerrungen bei der Schätzung mithilfe geeigneter statistischer Verfahren korrigiert werden (Europäische Kommission, 2016; 2018). Statistische Verfahren, die dafür infrage kommen, sind die Heckman-Methode, die Stratifizierung, die Clusteranalyse sowie die Extremwert-Methode.

- i) **Heckman-Methode:** Bei der Heckman-Methode wird die Wahrscheinlichkeit modelliert, dass ein Unternehmen einer Betriebsprüfung unterzogen wird. Auf Basis der für jedes Unternehmen separat berechneten Wahrscheinlichkeit wird die nach oben verzerrte Schätzung der Unternehmensteuerlücke korrigiert und somit die auf Basis der

¹⁵ Das IRS veröffentlicht beispielsweise jedes Jahr die Anzahl der geprüften Steuererklärungen, welche mathematische Fehler beinhalteten und wodurch diese entstehen: <https://view.officeapps.live.com/op/view.aspx?src=https%3A%2Fwww.irs.gov%2Fpub%2Firs-soi%2F20dbs03t23nm.xlsx&wdOrigin=BROWSELINK>

- geprüften Unternehmen erzielten Ergebnisse an die nicht geprüften Unternehmen angepasst.
- ii) **Stratifizierung:** Bei der Stratifizierung werden manuell Unternehmensgruppen identifiziert, die sich im Hinblick auf bestimmte Eigenschaften (Unternehmenstyp, Industrie, Grad an internationaler Verflechtung, etc.) ähneln. Der Methode liegt die Annahme zugrunde, dass sich innerhalb der Gruppen das Hinterziehungsverhalten der Unternehmen ähnelt. Für jede Gruppe wird anhand der Betriebsprüfungsergebnisse der geprüften Unternehmen die Steuerlücke errechnet. Die Summe der Gruppen-Steuerlücken ergibt die Unternehmensteuerlücke.
 - iii) **Clusteranalyse:** Die Clusteranalyse ähnelt der Stratifizierung, mit dem Unterschied, dass die Unternehmensgruppen anhand von Methoden maschinellen Lernens und nicht manuell gebildet werden.
 - iv) **Extremwert-Methode:** Die Extremwert-Methode wird häufig genutzt, wenn die Stichprobe an Unternehmen zu klein und heterogen ist, um die Ergebnisse mithilfe statistischer Verfahren auf die Grundgesamtheit zu übertragen (Bloomquist et al., 2014). Studien haben ergeben, dass sowohl die verdeckten Gewinne von Unternehmen als auch die Unternehmensgröße Pareto-verteilt sind (IRS, 2019; Axtell, 2001; Krishnaji, 1970). Da risikobasierte Betriebsprüfungen auf Unternehmen abzielen, die eine höhere Wahrscheinlichkeit aufweisen, Gewinne zu unterschlagen, können die Extremweltergebnisse der größten Unternehmen auf Basis der Pareto-Verteilung extrapoliert werden, um die Steuerlücke auf nationaler Ebene zu schätzen.

3.4.2 Ergebnisse

Risikobasierte Betriebsprüfungen werden häufig in Kombination mit randomisierten Prüfungen eingesetzt.

USA: Seit den 1960ern wählt die IRS jährlich eine große Anzahl an Unternehmen aus und unterzieht sie einer intensiven Prüfung. Für jede Position in der Steuererklärung wird die Angabe des Steuerpflichtigen mit der gemäß der Prüfung korrekten Angabe verglichen. Ausgehend davon wird die aggregierte Steuerlücke und ihre Zusammensetzung berechnet. Hierfür werden jedoch nicht nur die Unternehmensteuern geprüft, sondern auch Einkommens-, Umsatz-, Lohn- und Erbschaftssteuer. Der Großteil der Betriebsprüfungen erfolgt auf Basis einer randomisierten Auswahl, jedoch werden auch Unternehmen auf Basis des geschätzten Risikos ausgewählt (IRS, 2019). Das IRS benennt als die wichtigsten Gründe für das Auftreten einer Steuerlücke *non-filing*, *underreporting* und *underpayment*. *Non-filing* bezeichnet das Nichteinreichen einer Steuererklärung. Dieser Faktor wird jedoch nicht bei der Berechnung der Unternehmensteuerlücke berücksichtigt. *Underreporting* bezieht sich vor allem auf unterschlagene oder zu gering angegebene steuerpflichtige Einkünfte. *Underpayment* bezeichnet Steuerverbindlichkeiten, welche zeitgerecht gemeldet, aber nicht gezahlt wurden (IRS, 2019). Der auf *underreporting* zurückzuführende

Teil der Unternehmensteuerlücke wird separat für Klein-¹⁶ und Großunternehmen geschätzt. Für Kleinunternehmen wird ein ökonometrisches Modell zur Berücksichtigung des Selektionseffekts, der bei risikobasierten Betriebsprüfungen entsteht, genutzt, welches der Heckman-Methode ähnelt. Für Großunternehmen wird die Extremwert-Methode genutzt. Die Extremweltergebnisse der größten US-amerikanischen Unternehmen (Unternehmen mit einem Vermögen größer als 250 Millionen US-Dollar) werden anhand der Pareto-Verteilung angepasst, um das Ausmaß an *underreporting* für die Gesamtheit aller Großunternehmen angeben zu können. Beim Verfahren zur Berechnung des Ausmaßes an *underreporting* wird weder bei Klein- noch bei Großunternehmen eine Korrektur dafür vorgenommen, dass ein Teil der Steuerlücke unentdeckt bleibt, sodass die Ergebnisse als Untergrenze für die wahre Höhe der Steuerlücke interpretiert werden können (IRS, 2019).

Die aktuellste Hochrechnung aus dem Jahr 2019 schätzt die „Brutto“-Steuerlücke¹⁷ für den Zeitraum von 2011 bis 2013 auf durchschnittlich 441 Milliarden US-Dollar pro Jahr (18,3 Prozent der eigentlichen Steuereinnahmen). Davon entfallen 42 Milliarden US-Dollar auf die Unternehmensteuerlücke, wovon insgesamt 37 Milliarden US-Dollar auf unterschlagene Einkünfte (*underreporting*) zurückzuführen sind. Der Beitrag von Kleinunternehmen beträgt 11 Milliarden US-Dollar, der von Großunternehmen 26 Milliarden US-Dollar.¹⁸ Um die Netto-Unternehmensteuerlücke von 32 Milliarden US-Dollar zu ermitteln, werden von der Bruttolücke geleistete Nachzahlungen und verspätet eingegangene Zahlungen abgezogen. Auf Basis der Bruttosteuerlücke und der vor den Betriebsprüfungen eingekommenen Steuern wird zudem eine *Voluntary Compliance Rate* (VCR) berechnet, die dem Verhältnis der beiden Variablen entspricht (gezahlte Steuern als Anteil an der Bruttounternehmensteuerlücke). Die VCR misst, wie hoch der Anteil der „freiwillig“ gezahlten Steuern ist und kann als Maß für die Steuerehrlichkeit interpretiert werden (IRS, 2019).

Italien: Die italienische Steuerbehörde IRA (*Italian Revenue Administration*) führt seit 1998 jährlich risikobasierte Betriebsprüfungen bei Kleinunternehmen durch (*Studi di Settore*). Für die Korrektur des Selektionseffekts wird eine nachträgliche Stratifizierung angewendet. Hierfür werden die Steuerpflichtigen manuell in Gruppen aufgeteilt basierend auf Eigenschaften, die für die Auswahl der zu prüfenden Unternehmen von Bedeutung sind.

Kanada: Die kanadische Bundesbehörde CRA führt jährlich risikobasierte Betriebsprüfungen bei Unternehmen durch. Die Risikobewertung basiert auf dem vergangenen Verhalten der Unternehmen sowie der Art der Aktivitäten der Unternehmen. Für die Erfassung der Unternehmenssteuerlücke werden randomisierte Betriebsprüfungen bei mittelgroßen und Kleinunternehmen (KMU)

¹⁶ Unternehmen ohne Bilanzaufstellung oder mit einem Vermögen das kleiner ist als 10 Millionen US-Dollar.

¹⁷ Die Steuerlücke infolge von nicht rechtzeitig geleisteten und nicht freiwillig gezahlten Steuern.

¹⁸ Eine umfangreiche Übersicht ist: <https://www.irs.gov/pub/irs-pdf/p5365.pdf>

sowie, aufgrund ihrer geringen Anzahl¹⁹, risikobasierte Prüfungen bei Großunternehmen durchgeführt. Insgesamt werden 63 Großunternehmenscluster erstellt, für welche jeweils die Unternehmensteuerlücke separat errechnet wird. Die Summe der Cluster-Steuerlücken ergibt die Unternehmensteuerlücke basierend auf der Clusteranalyse. Die Steuerlücke für Großunternehmen liegt zwischen 1,7 und 2,9 Milliarden kanadische Dollar, was 4-7 Prozent der Unternehmensteuereinnahmen entspricht. Die Autoren der Studie räumen ein, dass Betriebsprüfungen teilweise unterschlagenes Einkommen nicht aufdecken und das Ergebnis somit als Untergrenze gesehen werden sollte (CRA, 2019).

Großbritannien und Australien: Wie oben erwähnt führen beide Länder neben randomisierten Betriebsprüfungen auch risikobasierte Prüfungen bei Mittel- und Großunternehmen durch. Für die Korrektur des Selektionseffekts wendet die ATO eine modifizierte Version der Extremwert-Methode an.²⁰ Die britische HMCR nutzt ebenfalls die Extremwert-Methode sowie eine Oberer-Grenzwert-Methode (HMCR, 2022).²¹

3.4.3 Bewertung

Im Hinblick auf die Schätzung der Unternehmensteuerlücke weist die Durchführung risikobasierter Betriebsprüfung gegenüber randomisierten Betriebsprüfungen einen gravierenden Nachteil auf. Eine risikobasierte Auswahl von Unternehmen für Betriebsprüfungen erschwert die Hochrechnung der auf Basis der Stichprobe ermittelten Steuerlücke auf die Grundgesamtheit aller inländischen Unternehmen. Anders als bei einer randomisierten Durchführung von Betriebsprüfungen kann bei risikobasierten Betriebsprüfungen nämlich nicht davon ausgegangen werden, dass die geprüften Unternehmen die Grundgesamtheit repräsentativ abbilden. Bei naiver Extrapolation der Stichprobenergebnisse wird die Unternehmensteuerlücke vermutlich überschätzt, da Unternehmen, bei denen das Risiko für das Auftreten einer Diskrepanz zwischen Steuerschuld und Steuerzahlung groß ist, mit größerer Wahrscheinlichkeit geprüft werden. Es existieren zwar verschiedene statistische Verfahren, um diese Verzerrung zu korrigieren; allerdings basieren die Verfahren auf Annahmen, deren Gültigkeit sich mitunter nur schwer oder überhaupt nicht überprüfen lassen.

Das risikobasierte Betriebsprüfungen trotzdem weit verbreitet sind, hat andere Gründe. Zum einen sind die (Netto-)Kosten im Vergleich zu randomisierten Betriebsprüfungen geringer, da Betriebsprüfungen basierend auf Risikoeinschätzungen in der Regel höhere Rückzahlungen abwerfen. Darüber hinaus erhöhen sie aus Perspektive der Unternehmen auch die Wahrscheinlichkeit der Aufdeckung von Steuerhinterziehung, wodurch sie einen größeren unmittelbaren Effekt auf

¹⁹ 1 Prozent aller Unternehmen fallen in die Gruppe der Großunternehmen. Sie tragen jedoch 54 Prozent zu den Unternehmensteuereinnahmen bei.

²⁰ <https://www.ato.gov.au/About-ATO/Research-and-statistics/In-detail/Tax-gap/Medium-business-income-tax-gap/?anchor=Methodology#Step1EstablishALogisticRegressionTrend>

²¹ Bei dieser Methode wird angenommen, dass nicht geprüfte Unternehmen im Schnitt eine ebenso hohe Steuerlücke aufweisen wie geprüfte Unternehmen.

die Schließung der Steuerlücke haben, als dies bei randomisierten Betriebsprüfungen der Fall ist. Schließlich erlauben risikobasierte Betriebsprüfungen einen effizienteren Ressourceneinsatz in Fällen, in denen eine Betriebsprüfung mit höherem Aufwand verbunden ist, wie beispielsweise bei Großunternehmen oder Selbständigen.

4 Perspektiven für die Steuerverwaltungspraxis

Eine randomisierte Auswahl für die Durchführung von Betriebsprüfungen stellt nicht nur das am besten geeignete Verfahren dar, um die Unternehmensteuerlücke sowie den Beitrag der einzelnen Komponenten, die ihre Höhe bestimmen, verlässlich zu schätzen. Die randomisierte Auswahl bietet darüber hinaus das Potenzial, die Effizienz der Betriebsprüfungspraxis zu erhöhen. Betriebsprüfungen sind mit einem hohen zeitlichen und personellen Aufwand verbunden. Die Ressourcen, die der Steuerverwaltung hierfür zur Verfügung stehen, sind jedoch begrenzt. Nicht selten sind die Steuerabteilungen von Großunternehmen den für die Betriebsprüfungen verantwortlichen Behörden im Hinblick auf die Ressourcenausstattung deutlich überlegen. Der daraus resultierende Rückstand der Steuerverwaltung ließe sich jedoch durch den Einsatz moderner Verfahren zur Erfassung und Auswertung von Daten – insbesondere solche, die unter dem Überbegriff Maschinelles Lernen zusammengefasst werden – reduzieren. Die auf Basis von randomisierten Betriebsprüfungen gewonnenen Informationen können dafür eine wichtige Grundlage bieten.

Unter dem Begriff Maschinelles Lernen (ML) versteht man allgemein Verfahren, bei denen sich mithilfe von Algorithmen Muster in Daten identifizieren lassen (Athey, 2018; Athey & Imbens, 2019). Ziel ML-basierter Verfahren ist in der Regel, auf Grundlage der identifizierten Muster Vorhersagen zu treffen oder Klassifizierungen vorzunehmen. ML-basierte Verfahren können beispielsweise herangezogen werden, um auf Basis der Realisationen mehrerer Variablen X_i (in der ML-Literatur Eigenschaften oder „Features“ genannt) die Ausprägung einer anderen Variable Y vorherzusagen. Zur Identifikation der Muster bzw. eines Modells, das der Vorhersage dient, wird ein sogenannter Trainingsdatensatz verwendet. Anschließend wird das Prognosemodell auf Beobachtungen abseits des Trainingsdatensatzes angewendet.

Die Informationen, die im Rahmen randomisierter Betriebsprüfungen gewonnen werden, erfüllen dabei zwei Funktionen. Erstens können sie für die Erstellung des Trainingsdatensatzes genutzt werden. Auf Basis des Trainingsdatensatzes lässt sich ein Modell identifizieren, das es der Steuerverwaltung erlaubt, für jedes beliebige Unternehmen das Risiko einzuschätzen, dass die Steuererklärung des Unternehmens fehlerhafte oder falsche Angaben enthält. Durch die Randomisierung ist sichergestellt, dass die in den Trainingsdaten identifizierten Muster repräsentativ sind für die inländischen Unternehmen und das Risikomodell entsprechend für sämtliche Unternehmen eine nützliche Risikoprognose liefert.

Zweitens erlaubt es die parallele Durchführung randomisierter Betriebsprüfungen und risikobasierter Betriebsprüfungen, das Risikomodell kontinuierlich anzupassen. Aus spieltheoretischer Perspektive lässt sich Steuervermeidung bzw. Steuerehrlichkeit als das Ergebnis einer strategischen Interaktion zwischen Unternehmen und der Steuerverwaltung interpretieren. In diesem Zusammenhang zeigt O’Doherty (2014) auf, dass die beste Strategie, um ein möglichst hohes Mehrergebnis zu erzielen und die Steuerehrlichkeit zu fördern, eine „lernende Steuerverwaltung“

ist, die ihr Risikomodell auf Basis neuer Informationen regelmäßig anpasst. Dahinter steckt der Gedanke, dass Unternehmen im Zeitablauf lernen, wie das Risikomodell der Steuerverwaltung funktioniert und daraufhin ihre Aktivitäten so anpassen, dass die Wahrscheinlichkeit, geprüft zu werden, möglichst gering wird. Durch die kontinuierliche Anpassung des Risikomodells kann die Steuerverwaltung dem entgegenwirken.

Die dritte Funktion der randomisierten Betriebsprüfung besteht darin, ein Gütemaß für das Risikomodell abzuleiten. Die Güte des Risikomodells lässt sich bestimmen, indem man das Mehrergebnis infolge von risikobasierten Betriebsprüfungen mit dem Mehrergebnis infolge randomisierter Betriebsprüfungen vergleicht. Ist der Informationsgehalt des Risikomodells positiv, so sollte es entsprechend zu einem höheren Mehrergebnis führen als im Fall einer zufallsbasierten Auswahl von Unternehmen für die Betriebsprüfung.

Entschließt sich die Steuerverwaltung in Zukunft für den Einsatz eines ML-basierten Risikomodells im Rahmen der Betriebsprüfung, könnte es außerdem von Interesse sein, zumindest in einer Übergangsphase parallel zum risikobasierten Auswahlverfahren und dem randomisierten Auswahlverfahren das aktuell angewandte Verfahren weiter laufen zu lassen. Dies würde einen direkten Vergleich ermöglichen zwischen der aktuellen Praxis und dem datengetriebenen, ML-basierten Risikomodell, wodurch sich – sofern sich dessen Überlegenheit bewahrheitet – die Akzeptanz des ML-basierten Verfahrens erhöhen ließe.

Als Vorbild für ein ML-basiertes Risikomodell kann die Praxis des britischen HMRC dienen. Für die Identifikation eines Risikomodells kombiniert das HMRC Informationen aus vergangenen Betriebsprüfungen mit Angaben aus den Steuererklärungen der geprüften Unternehmen. Das Risikomodell basiert auf einem sogenannten Entscheidungsbaum. Bei dieser Methode werden die Steuerpflichtigen in einem mehrstufigen Verfahren auf Basis eines Algorithmus in hierarchisch angeordnete Gruppen eingeordnet. Für jede dieser Gruppen wird dabei die Wahrscheinlichkeit ermittelt, dass eine Betriebsprüfung zu einer Steuernachzahlung führt. Jedes Unternehmen, das für das aktuelle Jahr eine Steuererklärung abgibt, wird anschließend auf Basis der in den Steuererklärungen enthaltenen Informationen (und ggf. weiterer, aus externen Datenquellen stammender Informationen) einer Gruppe zugeteilt. Aus der Gruppenzuteilung ergibt sich dann das Risiko.

5 Zusammenfassung und Schlussfolgerungen

Die Kurzexpertise stellt verschiedene Ansätze für die Schätzung der Unternehmensteuerlücke dar. In einer engen Definition versteht man unter der Steuerlücke die Differenz aus der gesetzlich geschuldeten Steuer und den tatsächlich gezahlten Steuern der steuerpflichtigen Unternehmen. Ursachen für das Auftreten einer Steuerlücke können sowohl Steuerhinterziehung als auch fehlende oder falsche Angaben in den Steuererklärungen sein, die auf Unkenntnisse oder Fehleinschätzungen der Steuerpflichtigen zurückzuführen sind. Für die Schätzung der Unternehmensteuerlücke gemäß der engen Definition werden üblicherweise Informationen aus Betriebsprüfungen herangezogen. Diese Praxis ist in vielen Ländern verbreitet und die Ergebnisse dieser Schätzungen werden von den nationalen Steuerbehörden in regelmäßigen Abständen veröffentlicht. Es gibt zwischen den Ländern allerdings Unterschiede im Hinblick darauf, nach welchen Kriterien Unternehmen für Betriebsprüfungen ausgewählt werden. Während in einigen Ländern die Durchführung von Betriebsprüfungen zu einem großen Teil randomisiert erfolgt, führen andere Länder hauptsächlich risikobasierte Betriebsprüfungen durch.

Geht es um die Schätzung der Steuerlücke gemäß der engen Definition, sind randomisierte Betriebsprüfungen von Vorteil. Die Randomisierung stellt nämlich sicher, dass die Stichprobe an Unternehmen, die einer Betriebsprüfung unterzogen wird, repräsentativ ist für die Grundgesamtheit aller inländischen Unternehmen. Die auf Basis der Stichprobe ermittelte Steuerlücke lässt sich somit einfach durch Extrapolation auf nationale Ebene hochrechnen. Erfolgt die Auswahl von Unternehmen für Betriebsprüfungen dagegen risikobasiert, ist die Repräsentativität nicht gewährleistet. Durch einfache Extrapolation würde man das Ausmaß der Unternehmensteuerlücke vermutlich überschätzen, da vor allem solche Unternehmen einer Betriebsprüfung unterzogen werden, bei denen die Wahrscheinlichkeit, dass die gezahlten Steuern geringer sind als die geschuldeten Steuern, besonders groß ist. Zwar existieren statistische Verfahren, mit deren Hilfe sich die entstehende Verzerrung korrigieren lässt; allerdings beruhen diese auf Annahmen, deren Gültigkeit nur schwer oder gar nicht überprüft werden kann.

Es gibt jedoch auch Gründe, die für die Durchführung risikobasierter Betriebsprüfungen sprechen. Der wohl wichtigste ist, dass risikobasierte Betriebsprüfungen eine höhere Wahrscheinlichkeit haben, besonders gravierende Fälle von Steuerhinterziehung aufzudecken, und erwartungsgemäß zu höheren Nachzahlungen führen. Wägt man die Vor- und Nachteile randomisierter und risikobasierter Betriebsprüfungen gegeneinander ab, spricht vieles dafür, beide Verfahren zu kombinieren. Durch eine Kombination ließe sich darüber hinaus die Effizienz risikobasierter Betriebsprüfungen verbessern.

Auf Basis der Informationen, die aus randomisierten Betriebsprüfungen gewonnen werden, ließe sich durch Methoden Maschinellen Lernens ein Risikomodell schätzen, das die Auswahl an Unternehmen für risikobasierte Betriebsprüfungen datengestützt anleitet. Ein solches Vorgehen hätte das Potenzial, den Zeit- und Ressourceneinsatz der Steuerverwaltung zu optimieren und

das Mehrergebnis zu erhöhen. Als Vorbild könnte hierbei die britische Steuerbehörde (britische *Her Majesty's Revenue and Customs*) dienen, die ein auf Maschinellem Lernen basiertes Risikomanagement seit einiger Zeit erfolgreich einsetzt.

In einer weiten Definition der Unternehmensteuerlücke umfasst diese zusätzlich noch den Verlust an Aufkommen aus Unternehmensteuern, der auf eine steuermotivierte Verlagerung von Gewinnen durch multinationale Unternehmen zurückzuführen ist. Für die Quantifizierung des Aufkommensverlusts kommen ein regressionsbasierter Ansatz oder ein Matching-Ansatz infrage. Die Trotz der weiten Verbreitung insbesondere des regressionsbasierten Ansatzes sollten dessen Ergebnisse mit Vorsicht interpretiert werden, da die Ergebnisse sehr sensitiv gegenüber der Spezifikation des zugrundeliegenden Regressionsmodells sein können. Eine Herausforderung in Bezug auf die Anwendung des regressionsbasierten Ansatzes besteht darin, dass umfangreiche Informationen über die Inlands- und Auslandsaktivitäten sämtlicher im Inland tätigen multinationalen Unternehmen vorhanden sein müssen, um die Höhe des Aufkommensverlusts infolge steuermotivierter Gewinnverlagerung zu schätzen.

Literaturverzeichnis

- Advani, A., Elming, W., & Shaw, J. (2017). The dynamic effects of tax audits. *The Review of Economics and Statistics*, 1-45.
- Athey, S. (2018). The impact of machine learning on economics. In: Goldfarb, A., Gans, J., & Agrawal, A: *The economics of artificial intelligence: An agenda*. University of Chicago Press, 507-547.
- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685-725.
- Axtell, R. L. (2001). Zipf distribution of US firm sizes. *Science*, 293(5536), 1818-1820.
- Beer, S., de Mooij, R., & Liu, L. (2020a). International corporate tax avoidance: A review of the channels, magnitudes, and blind spots. *Journal of Economic Surveys*, 34(3), 660-688.
- Beer, S., Kasper, M., Kirchler, E., & Erard, B. (2020b). Do audits deter or provoke future tax non-compliance? Evidence on self-employed taxpayers. *CESifo Economic Studies*, 66(3), 248-264.
- Bérgolo, M. L., Ceni, R., Cruces, G., Giacobasso, M., & Perez-Truglia, R. (2018a). Tax audits as scarecrows: Evidence from a large-scale field experiment. National Bureau of Economic Research Working Paper No. w23631.
- Bérgolo, M., Ceni, R., Cruces, G., Giacobasso, M., & Perez-Truglia, R. (2018b). Misperceptions about tax audits. *American Economic Review: Papers and Proceedings*, 108, 83-87.
- Bloomquist, K. M., Hamilton, S., & Pope, J. (2014). Estimating corporation income tax under-reporting using extreme values from operational audit data. *Fiscal Studies*, 35(4), 401-419.
- Bratta, B., Santomartino, V., & Acciari, P. (2021). Assessing profit shifting using country-by-country reports: A non-linear response to tax rate differentials. Dipartimento delle Finanze Working Paper 11.
- Clausing, K. A. (2009). Multinational firm tax avoidance and tax policy. *National Tax Journal*, LXII (4), 703-725.
- Clausing, K. A. (2011). The revenue effects of multinational firm income shifting. *Tax Notes*, March 28, 2011.

- CRA (2019). Tax gap and compliance results for the federal corporate income tax system. URL: <https://www.canada.ca/content/dam/cra-arc/corp-info/aboutcra/tax-gap/txgp2019-en.pdf>
- D'Agosto, E., Manzo, M., Modica, A., & Pisani, S. (2017). Tax audits and tax compliance: Evidence from Italy.
- DeBacker, J., Heim, B. T., Tran, A., & Yuskavage, A. (2015). Legal enforcement and corporate behavior: An analysis of tax aggressiveness after an audit. *The Journal of Law and Economics*, 58(2), 291-324.
- Dowd, T., Landefeld, P., & Moore, A. (2017). Profit shifting of U.S. multinationals. *Journal of Public Economics*, 148, 1-13.
- Egger, P., Eggert, W., & Winner, H. (2010). Saving taxes through foreign plant ownership. *Journal of International Economics*, 81(1), 99-108.
- Europäische Kommission (2016). The concept of tax gaps: Report on VAT gap estimation. Directorate-General for Taxation and Customs Union. URL: https://ec.europa.eu/taxation_customs/system/files/2016-09/tgpg_report_en.pdf.
- Europäische Kommission (2018). The concept of tax gaps: Report II, Corporate income tax gap estimation methodologies. Working Paper No. 73-2018. URL: <https://data.europa.eu/doi/10.2778/83206>
- Erard, B., & Feinstein, J. (2011). The individual income reporting gap: what we see and what we don't. In IRS-TPC Research Conference on New Perspectives in Tax Administration.
- Feinstein, J. S. (1999). Approaches for estimating noncompliance: Examples from Federal Taxation in the United States. *The Economic Journal*, 109, 360-369.
- Finke, K. (2013). Tax avoidance of German multinationals and implications for tax revenue: Evidence from a propensity score matching approach. mimeo.
- Fuest, C., Hugger, F., & Neumeier, F. (2022). Grenzüberschreitende Geschäftsbeziehungen innerhalb von Unternehmensgruppen – Ausmaß und Reformoptionen. ifo Studie im Auftrag des Bundesministeriums der Finanzen, München.
- Garcia-Bernado, J., & Janský, P. (2020). Profit shifting by multinational corporations worldwide: Evidence from country by country reporting. mimeo.
- Gemmell, N., & Hasseldine, J. (2012). The tax gap: a methodological review. Emerald Group Publishing Limited.

- Gemmell, N., & Ratto, M. (2012). Behavioral responses to taxpayer audits: Evidence from random taxpayer inquiries. *National Tax Journal*, 65(1), 33-57.
- Hansson, Å., Olofsdotter, K., & Thede, S. (2018). Do Swedish multinationals pay less in taxes than domestic firms? *The World Economy*, 41(2), 393-413.
- Heckemeyer, J. H., & Overesch, M. (2017). Multinational's profit response to tax differentials: Effect size and shifting channels. *Canadian Journal of Economics*, 50(4), 965-994.
- HMCR (2020). Non-detection multipliers for measuring tax gaps. Research and Analysis. URL: <https://www.gov.uk/government/publications/non-detection-multipliers-for-measuring-tax-gaps/non-detection-multipliers-for-measuring-tax-gaps>.
- HMCR (2021). Measuring tax gaps 2021 edition - Tax gap estimates for 2019 to 2020. Official statistics. URL: <https://www.gov.uk/government/statistics/measuring-tax-gaps/measuring-tax-gaps-2021-edition-tax-gap-estimates-for-2019-to-2020#corporation-tax-1>
- HMCR (2022). Measuring tax gaps 2021 edition - methodological annex. Official statistics. URL: <https://www.gov.uk/government/statistics/measuring-tax-gaps/measuring-tax-gaps-2021-edition-methodological-annex>.
- Hoopes, J. L., Mescall, D., & Pittman, J. A. (2012). Do IRS audits deter corporate tax avoidance? *The Accounting Review*, 87(5), 1603-1639.
- IRS (2019): Federal tax compliance research: Tax gap estimates for tax years 2011-2013. Washington, DC.
- Johansson, Å., Øystein, B. S., Sorbe, S., & Menon, C. (2017). Tax planning by multinational firms: Firm-level evidence from a cross-country database. OECD Economics Department Working Papers No. 1355.
- Kasper, M., & Alm, J. (2022). Audits, audit effectiveness, and post-audit tax compliance. *Journal of Economic Behavior & Organization*, 195, 87-102.
- Krishnaji, N. (1970). Characterization of the Pareto distribution through a model of underreported incomes. *Econometrica: Journal of the Econometric Society*, 251-255.
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651-692.
- O'Doherty, M. (2014). Thinking and learning in the tax evasion game. *Fiscal Studies*, 35(3), 297-339.

- OECD (2015). Measuring and monitoring BEPS. Action 11: 2015 final report. OECD Publishing, Paris.
- OECD (2017). Tax administration 2017: Comparative information on OECD and other advanced and emerging economies. OECD Publishing, Paris.
- Sarin, N., & Summers, L. H. (2019). Shrinking the tax gap: approaches and revenue potential. National Bureau of Economic Research Working Paper No. w26475.
- Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, 57(4), 904-54.
- Slemrod, J., Blumenthal, M., & Christian, C. (2001). Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics*, 79(3), 455-483.
- Søndergaard, J. (2018). Using random audit data to test the effect of certified accountancy on the tax compliance of small and medium enterprises in Denmark. *The IRS Research Bulletin*.
- Tørsløv, T.R., Wier, L.S., & Zucman, G. (2018). The missing profits of nations. National Bureau of Economic Research Working Paper No. w24701.
- Wissenschaftlicher Beirat beim Bundesministerium der Finanzen (2020). Notwendigkeit, Potenzial und Ansatzpunkte einer Verbesserung der Dateninfrastruktur für die Steuerpolitik. Gutachten 5/2020, Bundesministerium der Finanzen, Berlin.