



Effizienz oder Gerechtigkeit?
Ungleiche Einkommen,
ungleiche Vermögen und
die Theorie der optimalen
Besteuerung

Felix Bierbrauer





Effizienz oder Gerechtigkeit? Ungleiche Einkommen, ungleiche Vermögen und die Theorie der optimalen Besteuerung

Felix Bierbrauer

Februar 2016

Effizienz oder Gerechtigkeit? Ungleiche Einkommen, ungleiche Vermögen und die Theorie der optimalen Besteuerung*

Felix Bierbrauer**

Zusammenfassung

Felix Bierbrauer gibt einen Überblick der Theorie optimaler Steuern. Der Schwerpunkt liegt auf der Weiterentwicklung dieser Theorie seit der Jahrtausendwende. Er stellt aber auch die konzeptionellen Grundlagen dar sowie die klassischen Resultate zur optimalen Besteuerung von Arbeitseinkommen, Kapitaleinkommen und Konsumausgaben. Die Theorie hat in den letzten Jahren eine quantitative Dimension hinzugewonnen, die es beispielsweise erlaubt, die Einkommensteuerzahlungen von Spitzenverdienern im Status quo und unter einem optimal gestalteten Einkommensteuertarif miteinander zu vergleichen. Die jüngere Literatur hat außerdem geklärt, unter welchen Bedingungen die Beschäftigungsaufnahme von Geringverdienern subventioniert werden sollte. Der New-Dynamic-Public-Finance-Ansatz hat ein neues und wichtiges Argument entwickelt, das für eine Besteuerung von Kapitaleinkommen spricht. Weitere Themen der jüngeren Literatur sind der Wettbewerb verschiedener Regierungen um leistungsfähige und mobile Steuerzahler sowie die Interdependenz von optimaler Steuerpolitik und optimal gestalteten Staatsausgaben.

JEL: E21, H21, H24, H40

Schlüsselwörter: Optimale Besteuerung, Einkommensteuern, Verbrauchsteuern, Besteuerung von Kapitaleinkommen, Öffentliche Güter

* Ich danke Kollegen und Koautoren, die meinen Blick auf die Theorie optimaler Steuern in zahlreichen Gesprächen geprägt haben. Mein Dank gilt insbesondere Pierre Boyer, Emanuel Hansen, Martin Hellwig, Dominik Sachs und John Weymark. Ich danke auch dem Max Planck Institut zur Erforschung von Gemeinschaftsgütern in Bonn für seine Gastfreundschaft. Dort ist diese Arbeit entstanden. Für Kommentare zu diesem Manuskript danke ich Christoph Engel, Raphael Flore, Emanuel Hansen, Martin Hellwig, Alexander Morell, Dominik Sachs und Paul Schempp sowie dem Herausgeber Giacomo Corneo.

** Felix J. Bierbrauer, Center for Macroeconomic Research, Universität zu Köln, E-Mail: bierbrauer@wiso.uni-koeln.de

1 Einleitung

Wenn Politikerinnen und Politiker von SPD, Grünen oder Linken über die Steuerpolitik sprechen, betonen sie, dass sich in den letzten Jahren die Ungleichheit der Einkommen und Vermögen verschärft hat, dass das Einkommen eines Spitzenverdieners in der Wirtschaft ein Vielfaches des Einkommens eines „normalen“ Beschäftigten darstellt, dass die Kapitaleinkommen noch stärker konzentriert sind als die Arbeitseinkommen und die Zurückhaltung bei der Besteuerung von Vermögen und Erbschaften zur Ausbildung persistenter Ungleichheit beiträgt. Aufgabe der Steuerpolitik ist es, diese Ungleichheit zu korrigieren.

Konservative und Liberale haben eine Reihe von Einwänden: Zunächst ist da die Frage, ob es überhaupt wünschenswert ist, Ungleichheit zu korrigieren. Unterstellen wir, wir hätten einen leistungsfähigen Sozialstaat, der eine menschenwürdige Existenz für Alle garantiert. Dann wird es Ungleichheit unter all jenen geben, deren Einkommen und Vermögen einen höheren Lebensstandard erlaubt. Warum sollte man diese Form der Ungleichheit überhaupt korrigieren wollen? Und dann sind da noch die Kosten der Steuerpolitik: Einkommensteuern reduzieren die Arbeitsanreize, Mehrwertsteuer und Sozialabgaben machen Schwarzarbeit attraktiv, Steuern auf Kapitaleinkommen führen dazu, dass die Reichen das Land verlassen, und dann muss auch noch die gesamte Finanzverwaltung aus dem Steueraufkommen bezahlt werden. Wäre es da nicht einfacher, effizienter und gerechter, man hätte einen einheitlichen Steuersatz auf alle Konsumausgaben. Dann würde jeder, der in Deutschland Konsumausgaben tätigt, zur Finanzierung der Staatsausgaben herangezogen, selbst wenn er in Monaco wohnt oder ein großes Vermögen geerbt hat. Ein System mit einem einheitlichen Steuersatz und ohne Ausnahmeregelungen wäre außerdem gut für die Demokratie, weil es keine Rolle für Interessengruppen mehr gäbe, die versuchen, Ausnahmeregelungen im Steuerrecht durchzusetzen. Schließlich wäre es ein Gewinn an Gerechtigkeit, wenn „die Reichen“ nicht mehr durch die Beauftragung geschickter Berater ihr Einkommen klein rechnen konnten, während den abhängig Beschäftigten nichts anderes übrig bleibt als ihre Steuerlast zu akzeptieren.

Nun klingen diese Einwände im Prinzip gerechtfertigt, aber wenn sie so geballt vorgetragen werden, entsteht eben doch der Eindruck, dass Konservative und Liberale sich nicht besonders für die Ungleichheit der Einkommen und Vermögen interessieren und damit auch nicht für die damit einhergehende Ungleichheit an Lebenschancen. Dagegen wirkt die Betonung der Ungleichheit durch die politische Linke bisweilen ökonomisch naiv, so als sei die Summe der Einkommen und Vermögen in der Gesellschaft gegeben und könne ohne weitere Konsequenz von denen, die zu viel haben, zu denen, die zu wenig haben, umverteilt werden.

Wer Naivität in Bezug auf die Wirkungen der Steuerpolitik ebenso irritierend findet wie eine achselzuckende Reaktion auf zunehmende Ungleichheit, der ist bei der Theorie der optimalen Besteuerung an der richtigen Adresse. Sie liefert einen konzeptionellen Rahmen, der es erlaubt, sowohl über die Effizienzkosten als auch über die Wohlfahrtsgewinne umverteiler Steuern zu sprechen. Ein optimales Steuersystem ist demnach eines, bei dem die Effizienzkosten und die Wohlfahrtsgewinne in einem „vernünftigen“ Verhältnis stehen. Dieser Aufsatz

stellt zunächst die grundlegenden Prinzipien der Theorie optimaler Steuern vor und erläutert dann die Weiterentwicklungen, die diese Theorie in den letzten Jahren genommen hat.

Was kann man für praktische Fragen der Steuerpolitik dabei lernen? Hier sind ein paar Beispiele: Es wird klar, was die theoretische Fundierung der wohl gängigsten steuerpolitischen Empfehlungen – Bemessungsgrundlage verbreitern, Steuersätze senken, Schlupflöcher schließen! – von Professoren der Finanzwissenschaft ist. Es wird auch klar, dass eine vehemente Verfolgung von Verteilungszielen logisch nicht das gleiche ist wie die Forderung eines hohen Spitzensteuersatzes. Ob das so ist, hängt von den Eigenschaften der Einkommensverteilung ab. Denkbar ist auch, dass die Umverteilung von oben nach unten am besten funktioniert, wenn der Spitzensteuersatz sehr niedrig ist. Schließlich wird erklärt, warum es unter Verwendung der gängigen Modelle schwierig ist, eine überzeugende Begründung für hohe Steuern auf Kapitaleinkommen und Erbschaften zu finden. Die jüngere Literatur hat dagegen einige Argumente zusammengetragen, die für solche Steuern sprechen.

Dieser Text ist wie folgt aufgebaut: Abschnitt 2 stellt den konzeptionellen Rahmen der Optimalsteuertheorie dar. Es ist eine Theorie, die auf zwei grundlegenden Annahmen basiert. Erstens wird unterstellt, dass die Steuerpolitik dazu dient, Wohlfahrtsziele zu realisieren. Die Wohlfahrtsziele werden dabei so formuliert, dass ein Konflikt zwischen Effizienz- und Verteilungszielen entsteht. Weder vollkommene Effizienz noch vollkommene Gleichheit sind demnach wünschenswert. Die zweite Annahme ist, dass Steuerzahler in eigennütziger und opportunistischer Weise auf die durch die Steuerpolitik geschaffenen Verhaltensanreize reagieren. In Abschnitt 2 werden außerdem die klassischen Anwendungen dieser Theorie diskutiert, nämlich die optimale Gestaltung eines Einkommensteuertarifs und eines Systems von Verbrauchsteuern. Abschnitt 3 beginnt mit einer Liste wichtiger Fragen, die von den klassischen Analysen nicht beantwortet werden, und geht dann über zu einer Darstellung der Fortschritte, die es hierzu in den letzten Jahren gab. Die abschließenden Bemerkungen in Abschnitt 4 enthalten einen etwas spekulativen Ausblick auf die Forschungsthemen der kommenden Jahre, insbesondere die Vermutung, dass die Aufmerksamkeit der Forschung sich den Grundlagen zuwenden wird. Dazu gehören das für die Steuerzahler postulierte Verhaltensmodell, die Formalisierung von Verteilungszielen und die Beziehung zwischen Optimalsteuertheorie und politischer Ökonomie.

2 Die Grundlagen

Ehe wir uns den jüngeren Entwicklungen in der Theorie optimaler Steuern zuwenden, beschäftigen wir uns mit den Klassikern dieser Literatur. Zum einen muss man diese Referenzen kennen, um zu verstehen, was der Ausgangspunkt der Literatur der letzten zehn bis fünfzehn Jahre ist. Zum anderen ist die Auseinandersetzung mit den Klassikern nicht abgeschlossen. Wir werden sehen, dass auch jüngere Aufsätze deren Fragestellungen erneut aufgreifen.

2.1 Optimale Verbrauchsteuern: Ramsey (1927) und Boiteux (1956)

Die Theorie der optimalen Besteuerung beginnt mit einer Arbeit von Frank Ramsey aus dem Jahr 1927. Ramsey betrachtet einen Staat, der ein vorgegebenes Einnahmeziel hat. Die Bestimmung der Staatsausgaben ist nicht Teil der Analyse¹. Der Staat kann alle Güter und Leistungen besteuern und die Steuern dabei beliebig differenzieren. Es kann einen eigenen Steuersatz für Apfelsinen geben, einen anderen für Windeln und wieder einen anderen für Champagner. Nicht nur Konsumausgaben unterliegen der Besteuerung, auch Arbeitseinkommen kann, wiederum mit einem eigenen Steuersatz, besteuert werden. Ramsey stellt nun die Frage, wie die optimale Kombination dieser Steuersätze aussieht. Ramseys Vorgehen ist daher nicht vereinbar mit der Vorstellung, dass Verbrauchsteuersystem solle in dem Sinne neutral sein, dass alle Konsumgüter mit dem gleichen Steuersatz belegt werden. Außerdem unterstellt es, dass keine vom Verbrauch unabhängigen Kopf- oder Pauschalsteuern zur Verfügung stehen, um Staatseinnahmen zu generieren. Die Bedeutung dieser Annahme wird weiter unten ausführlicher diskutiert.

Die zentrale Innovation in Ramseys Ansatz ist, dass er dabei nicht ein von der Steuerpolitik unabhängiges Verhalten unterstellt. Vielmehr berücksichtigt er, dass die Nachfrage nach Apfelsinen, Windeln oder Champagner und auch das Arbeitsangebot auf die Steuerpolitik reagieren werden. Es geht also um eine optimale Steuerpolitik, die die Verhaltensreaktion der Steuerzahler berücksichtigt. Das zentrale Resultat von Ramseys Analyse ist als inverse Elastizitätenregel bekannt geworden. Es sollten diejenigen Güter hoher besteuert werden, bei denen mit geringeren Ausweichreaktionen der Steuerzahler zu rechnen ist. Wenn also die Nachfrage nach Apfelsinen stärker auf die Besteuerung reagiert als die Nachfrage nach Windeln, dann sieht ein optimales Steuersystem höhere Steuern auf Windeln und niedrigere Steuern auf Apfelsinen vor.

Eine alternative Interpretation des Ramsey-Problems stammt von Boiteux (1956). Boiteux stellt sich einen natürlichen Monopolisten vor, zum Beispiel einen Energieversorger, der hohe fixe Kosten decken muss, etwa die Kosten des Leitungsnetzes. Der Energieversorger stellt unterschiedliche Leistungen für unterschiedliche Kundengruppen, etwa Privatkunden und Geschäftskunden, bereit und muss entscheiden, welche dieser Leistungen welchen Deckungsbeitrag zu den fixen Kosten erbringen soll. Auch Boiteux leitet eine inverse Elastizitätenregel her: Der Deckungsbeitrag einer Leistung soll umso höher sein, je niedriger die Elastizität der Nachfrage.

Ausweichreaktionen sind schlecht. Warum ist ein optimales Steuersystem dadurch charakterisiert, dass Ausweichreaktionen minimiert werden? Eine Steuer auf Apfelsinen reduziert die Nachfrage, und so bleiben Apfelsinen unproduziert, für die es eigentlich eine Zahlungsbereitschaft gäbe. Diese nicht realisierten Tauschgewinne zwischen Apfelsinenbauern und Haushalten stellen in Ramseys Rahmen die Wohlfahrtskosten der Besteuerung dar. Ein optimales

1 Wir werden weiter unten über jüngere Arbeiten sprechen, die simultan ein optimales Ausgabenniveau und eine optimale Steuerpolitik bestimmen.

Steuersystem minimiert diese Wohlfahrtskosten und besteuert also diejenigen Güter umso stärker, bei denen Ausweichreaktionen gering ausfallen².

Ramseys formale Analyse stellt eine spezifische Form von Ausweichreaktionen ins Zentrum, nämlich die Verlagerung der Nachfrage von stärker besteuerten zu weniger besteuerten Gütern. Die grundsätzliche Einsicht, dass Steuern zu Ausweichreaktionen führen und dass diese zu Einbußen an Effizienz führen, gilt aber sehr viel allgemeiner. Wenn die Einführung einer Fenstersteuer dazu führt, dass Fenster zugemauert werden, ist das ebenso ineffizient wie das Ausrüsten von Supermärkten mit Computerterminals, die es möglich machen, offline erworbene Waren als Ergebnis einer steuerbegünstigten Onlinetransaktion zu deklarieren.

Viele kleine Verzerrungen sind besser als wenige große. Eine Besteuerung nach der inversen Elastizitätenregel erfordert insbesondere, dass alle Güter besteuert werden. Kein Gut ist von der Besteuerung ausgenommen. Allerdings variiert der Beitrag verschiedener Güter zum Erreichen des staatlichen Einnahmeziels mit der Elastizität der Nachfrage. Dieses Resultat rührt daher, dass die Wohlfahrtverluste aus einer Erhöhung des Steuersatzes umso höher sind, je höher der Steuersatz in der Ausgangslage ist. Eine geringfügige Erhöhung der Steuer auf Apfelsinen, ausgehend von einem Steuersatz von Null, hat vernachlässigbare Wohlfahrtskosten, nicht jedoch ausgehend von einem Steuersatz von 20 Prozent. Führt man ausgehend von einem Steuersatz von Null eine 1 prozentige Steuer auf Apfelsinenverkäufe ein, werden einige Käufer sich fortan gegen Apfelsinen entscheiden, nämlich diejenigen, deren Zahlungsbereitschaft zuvor knapp über den Kosten der Apfelsinenproduktion gelegen hat. Wenn aber die Zahlungsbereitschaft nur knapp über den Kosten liegt, sind die Tauschgewinne klein. Wenn die verdrängten Tauschgewinne klein sind, dann sind es auch die Wohlfahrtskosten der Besteuerung. Ausgehend von einem Steuersatz von 20 Prozent verdrängt man durch eine Steuererhöhung aber Tauschgewinne, bei denen die Zahlungsbereitschaft deutlich über den Kosten liegt, und daher sind auch die Wohlfahrtsverluste nicht mehr klein.

Eine Anwendung dieser Logik führt in konkreten steuerpolitischen Diskussionen zu der Empfehlung, Schlupflöcher zu schließen, die Bemessungsgrundlage zu verbreitern und die Steuersätze zu senken. Wenn in der Ausgangslage nur eine Steuer auf Apfelsinen erhoben wird, um das staatliche Einnahmeziel zu erreichen, dann können die Wohlfahrtskosten der Besteuerung gesenkt werden, wenn die Steuern auf Apfelsinen gesenkt und die Steuern auf Windeln und Champagner erhöht werden.

Verteilungsaspekte. In diesen elementaren Überlegungen sind Verteilungsaspekte nicht berücksichtigt. Unterstellen wir, dass Windeln überwiegend von Familien mit durchschnittli-

2 In quantitativen Analysen werden diese entgangenen Tauschgewinne häufig in Prozent des Bruttoinlandsprodukts ausgewiesen. Alternativ werden die Wohlfahrtskosten gelegentlich auch zu den erzielten Einnahmen ins Verhältnis gesetzt, so dass man ein Maß für die Wohlfahrtskosten erhält, die mit einem Euro Staatseinnahmen einhergehen. Allerdings sind diese Maße für sich genommen nicht aussagekräftig, weil den Kosten der Einnahmeerzielung nicht der Nutzen aus der staatlichen Bereitstellung von Gütern gegenübergestellt wird. Hier zeigen sich die Grenzen einer Analyse, die nur die Einnahmeseite des Staates betrachtet.

chem Einkommen gekauft werden und Champagner überwiegend von Yuppies, die über viel Geld verfügen. Unterstellen wir weiter, dass die Nachfrage nach Windeln weniger elastisch ist als die Nachfrage nach Champagner. Schließlich kann man sich ohne Windeln nicht wirklich helfen, während man statt Champagner ja auch Prosecco trinken konnte. Dann wurde die Anwendung der inversen Elastizitätenregel dazu führen, dass Familien stärker belastet werden als Yuppies, was eine problematische Verteilungswirkung wäre. Um diese Schlussfolgerung zu vermeiden, hat Diamond (1975) Ramseys Analyse um Verteilungsaspekte erweitert. Deren Berücksichtigung kann es wünschenswert machen, von der inversen Elastizitätenregel abzuweichen. Die grundsätzliche Logik bleibt aber erhalten. Für Güter, die von allen Haushalten gleichmäßig konsumiert werden, gilt die inverse Elastizitätenregel weiter. Auch gilt weiter, dass alle Güter einen Beitrag zum Erreichen des staatlichen Einnahmeziels leisten sollen.

Pragmatische und Fundamentale Ansätze. In der Theorie optimaler Steuern gibt es seit jeher einen Gegensatz zwischen pragmatischen Modellierungsansätzen und „tiefen“, theoretisch vollständig begründeten Ansätzen. Diese Unterscheidung wird uns in diesem Aufsatz immer wieder begegnen. Die Pragmatiker möchten sich nur mit „realistischen“ Politikempfehlungen beschäftigen. Die Alternative dazu stellt ein allokationstheoretisch fundierter Ansatz dar. Dabei wird nach einer optimalen Steuerpolitik in der Menge der Politiken gesucht, die mit der Ressourcenausstattung und der Informationsstruktur der Ökonomie vereinbar sind. Häufig führt dies zu unterschiedlichen Ergebnissen. Am Beispiel des Ramsey-Modells, das zum pragmatischen Lager zählt, lässt sich das gut illustrieren.

Unterstellen wir, dass es Haushalte gibt, die sich im Einkommen unterscheiden oder auch in ihrer Fähigkeit am Arbeitsmarkt hohe Einkommen zu generieren. Unterstellen wir weiterhin, dass die Regierung diesen Haushaltstyp beobachten kann. Ein optimales Steuersystem à la Ramsey geht mit Effizienzeinbußen einher. Das wäre aber gar nicht nötig. Man könnte stattdessen Steuern erheben, die von den Konsumausgaben oder dem Arbeitsangebot der Haushalte unabhängig sind. Um Verteilungsaspekte zu berücksichtigen, könnten diese Steuern für einkommensstarke Haushaltstypen höher sein als für einkommensschwache. Auch eine solche Steuerpolitik würde das Einnahmeziel erreichen, und auch eine solche Steuerpolitik hätte eine Wirkung auf das Verhalten der Haushalte. Diese Wirkung wäre aber auf Einkommenseffekte beschränkt, also auf eine Veränderung der Nachfrage nach Gütern und des Arbeitsangebots aufgrund von Veränderungen im verfügbaren Einkommen. Die Wohlfahrtskosten der Besteuerung resultieren aber aus Substitutionseffekten, also Veränderungen im Verhalten, die getrieben sind durch die Veränderung relativer Preise. Solche Substitutionseffekte treten bei einer verhaltensunabhängigen Form der Besteuerung nicht auf.

Die Betrachtung von verhaltensunabhängigen Kopf- oder Pauschalsteuern liegt dem zweiten Hauptsatz der Wohlfahrtstheorie zugrunde. Dieser Satz besagt, dass es möglich ist, Umverteilungsziele und Einnahmeziele des Staates so zu realisieren, dass es nicht zu Effizienzeinbußen kommt. Aus dieser Sicht wirkt der Ramsey-Ansatz künstlich, weil er auf einer Optimierung in einer Klasse suboptimaler Besteuerungsinstrumente beruht. Ein Pragmatiker würde dagegen

einwenden, dass Kopfsteuern im politischen Prozess typischerweise nicht betrachtet werden und es sich deshalb auch nicht lohnt, sich eingehender damit zu beschäftigen.³

2.2 Optimale Einkommensbesteuerung: Mirrlees (1971)

Die Arbeit von Mirrlees (1971) markiert in mehrfacher Hinsicht eine Wende in der Theorie optimaler Steuern. Erstens beschreibt sie ein Verfahren zur Charakterisierung optimaler nicht-linearer Steuern. Damit können erstmals Fragen des Typs „Wie progressiv sollte der Einkommensteuertarif sein?“ formal analysiert werden. Ramseys Analyse war auf lineare Steuern beschränkt, so dass es nicht möglich war, die Frage zu stellen, ob der Grenzsteuersatz bei einem Einkommen von 50.000 Euro ein anderer sein sollte als bei einem Einkommen von 75.000 Euro. Zweitens kommt es zu einer Versöhnung zwischen dem fundamentalen und dem pragmatischen Ansatz in der Analyse optimaler Steuerpolitik. Die Abwägung von Effizienz- und Verteilungszielen im Mirrlees-Modell hat eine „tiefe“ alloktionstheoretische Fundierung, die zugleich eine Rechtfertigung für die Verwendung verzerrender Steuern liefert.

Nicht-lineare Besteuerung. Mirrlees (1971) betrachtet ein Modell mit Individuen, die sich in ihren produktiven Fähigkeiten unterscheiden. Es gibt keine weitere Quelle von Heterogenität. Manche Individuen sind produktiv und können auf dem Arbeitsmarkt hohe Einkommen erzielen, weniger produktive Individuen erzielen bei gleichem Einsatz ein sehr viel geringeres Einkommen. Der Staat kann durch Einsatz einer Einkommensteuer die so erzielten Einkommen umverteilen. Es gibt keine weitere Notwendigkeit der Einnahmeerzielung. Die bei manchen Haushalten erzielten Steuereinnahmen werden vollständig an die anderen Haushalte ausgezahlt. Die Einkommensteuerzahlungen der ersten Gruppe sind damit positiv, die der zweiten negativ. In einem Modell mit gut verdienenden Ingenieurinnen und schlechter verdienenden Hausmeistern wäre also die Frage, welcher Anteil des Einkommens der Ingenieurinnen über das Steuersystem an die Hausmeister umverteilt werden sollte.

Formal ist das Problem die Bestimmung eines Steuertarifs T , der jedem möglichen Vorsteuereinkommen y eine Steuerzahlung $T(y)$ zuordnet, so dass das Nachsteuereinkommen oder die Konsumausgaben c eines Haushaltes sich als $c=y-T(y)$ ergeben. Die Steuerzahlung kann mit dem Einkommen variieren, ebenso kann der Grenzsteuersatz $T'(y)$ vom Einkommen abhängen.⁴ Zu klären ist nun, welche Eigenschaften ein optimal gewählter Einkommensteuertarif hat. Sollte er progressiv sein? Wie hoch sollte der Spitzensteuersatz sein? Wie hoch sollte der Eingangssteuersatz sein? Wie hoch sollten die Transfers für einkommensschwache Haushalte ausfallen?

3 Die Ausnahme ist die von Margaret Thatchers Regierung in England eingeführte Poll Tax, die allerdings äußerst unpopulär war und schnell wieder abgeschafft wurde.

4 Der Grenzsteuersatz gibt an, welcher Anteil eines zusätzlichen Euros an Vorsteuereinkommen umverteilt wird. Bei einem progressiven Einkommensteuertarif steigt dieser Anteil mit dem Vorsteuereinkommen, so dass eine sehr gut verdienende Ingenieurin, die ihr Vorsteuereinkommen um einen Euro steigert, von diesem Euro weniger behalten darf als eine nur ordentlich verdienende Ingenieurin

Ziel ist die Maximierung eines Wohlfahrtsmaßes, das so gewählt ist, dass Einkommensumverteilung prinzipiell wünschenswert ist. Die Wohlfahrt steigt, wenn die niedrigen Nachsteuer-einkommen der Hausmeister sich auf die hohen Nachsteuereinkommen der Ingenieurinnen zubewegen. Könnte man ohne Effizienzeinbußen umverteilen, dann würde man folglich die Besteuerung der Ingenieurinnen und die Transfers zu den Hausmeistern so gestalten, dass am Ende beide die gleichen Konsummöglichkeiten haben.⁵ Wie schon Ramsey (1927), bezieht Mirrlees aber die Verhaltensreaktion der Haushalte ein. Wenn die Ingenieurinnen mehr von ihrem Arbeitseinkommen abgeben müssen, verschlechtern sich ihre Anreize zur Einkommenserzielung.⁶ Damit müssen die Wohlfahrtsgewinne aus zusätzlicher Umverteilung zu den Effizienzkosten der Besteuerung in Beziehung gesetzt werden. Ein optimaler Einkommensteuertarif im Sinne von Mirrlees (1971) ist demnach eine optimale Reaktion auf den Konflikt zwischen Effizienz- und Verteilungszielen.

Die Charakterisierung eines optimalen Einkommensteuertarifs in diesem Rahmen ist das Thema zahlreicher Aufsätze.⁷ Die zentralen Resultate dieser Literatur lauten wie folgt: Erstens, die Einkommensteuerzahlungen der Geringverdiener sind negativ. Sie bekommen Transferzahlungen, so dass ihr Nachsteuereinkommen größer ist als ihr Vorsteuereinkommen. Hier zeigt sich das Umverteilungsziel. Zweitens, je besser die Fähigkeiten eines Haushalts zur Einkommenserzielung sind, desto reicher sollte er unter einem optimalen Einkommensteuertarif sein. Produktivere Haushalte haben also höhere Einkommen, sowohl vor als auch nach Besteuerung. Dieses Resultat ist auf die Effizienzkosten der Besteuerung zurückzuführen. Das Herstellen gleicher Konsummöglichkeiten für alle ist zwar möglich, aber nicht wünschenswert. Drittens, optimale Grenzsteuersätze sind niemals negativ. Sie sind positiv für alle Steuerzahler mit Ausnahme des Spitzenverdieners. Dessen Grenzsteuersatz ist Null.

Das zuletzt genannte Resultat ist mit Blick auf die steuerpolitische Diskussion bemerkenswert: Trotz des Wunschs, Einkommen umzuverteilen, sollen die Reichsten einen Spitzensteuersatz von Null haben. Betrachten wir als Gedankenexperiment eine Reform, die den Steuersatz auf gehobene Einkommen, etwa solche, die zwischen 75.000 und 76.000 Euro liegen, um 1 Prozent erhöht. Das verschlechtert die Arbeitsanreize für all jene, die ein Einkommen in genau diesem Bereich erzielen. Wenn sie mehr arbeiten, haben sie etwas weniger in der Tasche als in der Ausgangssituation. Es bringt aber zusätzliche Steuereinnahmen von all jenen, die mehr als 76.000 Euro verdienen. Sie zahlen 10 Euro mehr im Jahr. Die Gesamteinnahmen, die so entstehen, sind 10 Euro multipliziert mit der Zahl der Haushalte, deren Einkommen jenseits der 76.000 Euro liegt. Diese Einnahmen können umverteilt werden und ermöglichen eine Wohlfahrtssteigerung. Wie verändert sich dieses Kalkül, wenn wir überlegen, den Spitzenverdiener höher zu besteuern? Wir haben wieder Effizienzkosten durch eine Verschlechterung

5 Etwas genauer: Ingenieurinnen und Hausmeister sollten den gleichen Grenznutzen im Konsum haben. Bei einer additiv separablen Nutzenfunktion und konkavem Konsumnutzen ist das äquivalent zu einer Angleichung der Konsumniveaus.

6 Wenn die Hausmeister Transfers bekommen, wird sich das ebenfalls auf ihr Arbeitsangebot auswirken. Das ist dann aber ein aus Effizienz-sicht unproblematischer Einkommenseffekt.

7 Eine Systematisierung der verschiedenen Ansätze findet sich in Hellwig (2007a).

rung der Anreize für den Spitzenverdiener. Dem steht nun aber kein Wohlfahrtsgewinn mehr gegenüber: Da es niemanden gibt, der mehr verdient als der Spitzenverdiener, ist der Einnahmegewinn 10 Euro multipliziert mit Null, was zu einem Einnahmegewinn von Null Euro führt. Das ist ein schlechter Deal: Positive Grenzsteuersätze beim Spitzeneinkommen bedeuten Effizienzkosten ohne Wohlfahrtsgewinne. Dies bedeutet nicht, dass der Spitzenverdiener keine Steuern zahlt. Die positiven Grenzsteuersätze bei allen niedrigeren Einkommen führen zu einer positiven Steuerzahlung des Spitzenverdieners.⁸ Aber den letzten Euro, den er dazu verdient, soll er unter einem optimalen Einkommensteuertarif behalten dürfen. Die Auseinandersetzung mit der Relevanz dieses Resultats für praktische Fragen der Steuerpolitik ist ein Thema der jüngeren Literatur, siehe Diamond (1998), Saez (2001), Mankiw et al. (2009), und Diamond und Saez (2011). Dazu weiter unten mehr.

Bemerkenswert ist auch das Resultat, wonach negative Grenzsteuersätze nicht mit einem optimalen Einkommensteuertarif vereinbar sind. Dies schließt eine Subventionierung der Beschäftigungsaufnahme für die Bezieher niedriger Einkommen aus. Auch dieses Resultat hat die jüngere Literatur beschäftigt, und auch auf diese Diskussion werden wir noch genauer eingehen. Die Ablehnung subventionierter Beschäftigung ergibt sich wieder aus dem Zusammenspiel von Effizienzkosten und Wohlfahrtsgewinnen. Betrachten wir eine Ausgangssituation mit einem negativen Grenzsteuersatz bei niedrigen Einkommen: Negative Grenzsteuern führen zu Effizienzkosten, da Geringverdiener übermäßige Arbeitsanreize haben. Es ist daher möglich, die Subventionierung abzubauen ohne die Geringverdiener schlechter zu stellen. Zugleich wird es leichter, bei denen Steuern einzusammeln, die höhere Einkommen erzielen. Ein Abbau negativer Grenzsteuersätze geht also sowohl mit positiven Effizienzwirkungen als auch mit positiven Wohlfahrtswirkungen einher.

Allokationstheoretische Begründung. Betrachten wir wieder eine Welt, in der es Ingenieurinnen mit guten Verdienstmöglichkeiten gibt und Hausmeister mit schlechteren Verdienstmöglichkeiten. Eine verhaltensunabhängige Form der Besteuerung würde erfordern, dass man jeder Steuerzahlerin und jedem Steuerzahler ansehen kann, ob er zur privilegierten Gruppe gehört. Wer dazu gehört, muss Steuern zahlen, unabhängig davon, wie viel Einkommen tatsächlich erzielt wird. Die Einnahmen werden dann an die weniger Privilegierten ausgezahlt, wiederum unabhängig davon, was diese wirklich verdienen. Könnte man den Menschen ihre Verdienstmöglichkeiten an der Nasenspitze ablesen, wäre diese Form der Besteuerung einem Steuersystem à la Mirrlees überlegen, bei dem die Steuerzahlung vom tatsächlich erzielten Einkommen abhängt. Wäre die Steuerbasis das potentiell erzielbare Einkommen, käme man ohne die Effizienzkosten aus, die sich aus der Verschlechterung der Arbeitsanreize ergeben.

8 Die Steuerschuld des Spitzenverdieners ergibt sich aus der Summe der Grenzsteuersätze auf alle niedrigeren Einkommensniveaus. Formal gilt: $T(y_{\max}) = T(0) + \int_0^{y_{\max}} T'(y) dy$, wobei y_{\max} das Einkommen des Spitzenverdieners bezeichnet und $-T(0)$ den Transfer an Individuen ohne Einkommen. Aus dieser Formel wird ersichtlich, dass auch die Bezieher hoher Einkommen profitieren, wenn die Steuersätze im Bereich niedriger Einkommen gesenkt werden.

Diese Kritik verfängt aber nur unter der Annahme, dass die Verdienstmöglichkeiten eines Haushalts eine öffentlich verfügbare Information darstellen. Wenn jeder Steuerzahler selbst am besten einschätzen kann, welche Verdienstmöglichkeiten sich ihm bieten, funktioniert ein derartiges System der Besteuerung nicht. Im Jargon der ökonomischen Theorie spricht man von „privater Information über produktive Fähigkeiten“. Die Frage nach einem optimalen umverteilenden Steuersystem unter Bedingungen privater Information führt uns zurück zum Ansatz von Mirrlees: Das sogenannte Besteuerungsprinzip (Taxation Principle im Englischen) besagt, dass es keinen Allokationsmechanismus gibt, der ein besseres Ergebnis herbeiführen kann als ein optimal gestaltetes Einkommensteuersystem.⁹ Es bleibt einem demnach nichts anderes übrig als das tatsächlich realisierte Einkommen zur Grundlage der Besteuerung zu machen. Das Besteuerungsprinzip liefert damit die theoretische Rechtfertigung für die Verwendung von Einkommensteuern, die vom tatsächlich erzielten Markteinkommen abhängen.

2.3 Die gemeinsame Betrachtung von Verbrauch- und Einkommensteuern: Atkinson und Stiglitz (1976)

Inwiefern kann die Rechtfertigung verzerrender Steuern auch auf Verbrauchsteuern ausgedehnt werden? Atkinson und Stiglitz (1976) betrachten eine Erweiterung des Modells von Mirrlees: Haushalte entscheiden weiterhin darüber, wie viel Aufwand sie treiben, um Einkommen zu erzielen. Es gibt nun aber wie im Modell von Ramsey mehrere Verwendungsmöglichkeiten für das Nachsteuereinkommen. Man kann Windeln, Champagner oder Apfelsinen kaufen. Macht es nun Sinn, die Einkommensteuer um differenzierende Verbrauchsteuersätze, also jeweils eigene Steuersätze für Windeln, Champagner oder Apfelsinen, zu ergänzen? Die Antwort von Atkinson und Stiglitz lautet „nein“. Wenn Haushalte sich nur in ihren produktiven Fähigkeiten unterscheiden und ihre Zahlungsbereitschaften für Windeln, Champagner und Apfelsinen von diesen produktiven Fähigkeiten unabhängig sind, dann lassen sich die Ergebnisse der Steuerpolitik durch differenzierende Verbrauchsteuersätze nicht verbessern.¹⁰ Eine Besteuerung etwa gemäß der inversen Elastizitätenregel ist demnach nicht angezeigt. Ein einheitlicher Verbrauchsteuersatz für alle Güter, etwa ein Mehrwertsteuersystem ohne ermäßigte Steuersätze für bestimmte Güter, ist dagegen mit einem optimalen Steuersystem vereinbar. Ein solches Mehrwertsteuersystem wird aber auch nicht zwingend benötigt. Ein einheitlicher Steuersatz von Null ist ebenfalls mit einem optimalen Steuersystem kompatibel.

Betrachten wir folgende Ausgangssituation: Ein optimales umverteilendes Einkommensteuersystem ist installiert worden, und alle Verbrauchsteuersätze sind gleich Null. Könnte man die Ergebnisse verbessern, wenn man nun eine Steuer auf Champagner einführt? Eine solche Steuer ist, wie weiter oben diskutiert, mit Effizienzkosten verbunden, weil mögliche Tauschgewinne zwischen Champagnerproduzenten und Champagnertrinkern nicht realisiert werden. Sie hätte daher nur einen Sinn, wenn sie die Verschlechterung der Arbeitsanreize, die das op-

9 Formale Darstellungen dieses Resultats finden sich in Hammond (1979), Stiglitz (1982) Guesnerie (1995) und Bierbrauer (2011a).

10 Elementare Beweise dieses Resultats finden sich in Kaplow (2006) und Laroque (2005).

timale Einkommensteuersystem unweigerlich mit sich bringt, konterkarieren könnte. Wenn alle Steuerzahler die gleiche Zahlungsbereitschaft für Champagner haben, dann lässt sich so keine Verbesserung erzielen. Die Champagnersteuer trifft dann alle gleichmäßig, unabhängig davon, wie einkommensstark sie sind, und es ist nichts gewonnen. Wenn aber „die Reichen“ ein besonderes Verhältnis zum Champagner haben, sieht die Sache anders aus. Unterstellen wir, dass „die Reichen“ auch dann noch überdurchschnittlich viel Champagner trinken würden, wenn widrige Umstände sie dazu zwängen, mit dem Einkommen eines Normalverdieners über die Runden zu kommen - einen gewissen Stil ist man sich schließlich schuldig. Die echten Normalverdiener hätten diese Neigung nicht. Wieso Champagner trinken, wenn man auch ein ordentliches Bier bekommen kann? Wenn die Welt so wäre, gäbe es eine Rolle für die Champagnersteuer. Die Einkommensteuer verschlechtert die Anreize „der Reichen“, sich bei der Einkommenserzielung besonders anzustrengen. Sie reduzieren also ihr Einkommen relativ zu einer Laissez-Faire-Situation ohne Einkommensbesteuerung. Eine geschickt gestaltete Champagnersteuer würde genau dann zuschlagen, wenn jemand mit einem derart reduzierten Einkommen Champagner kaufen will. Die Reduktion des Arbeitseinsatzes wird weniger attraktiv, wenn dann auch noch der Champagner teurer wird.¹¹

Ob die Nachfrage nach bestimmten Gütern gleichmäßig in der Bevölkerung verteilt ist, ist eine empirische Frage. Aber selbst wenn diese Frage so beantwortet wird, dass differenzierende Verbrauchsteuersätze Bestandteil einer wohlfahrtsmaximierenden Steuerpolitik wären, stellt sich die Frage nach der Relevanz dieser Beobachtung für die praktische Steuerpolitik. Differenzierende Verbrauchsteuersätze werfen insbesondere die Frage auf, ob Präferenzunterschiede, hier zwischen Reichen und Normalverdienern, zum Anlass genommen werden sollten, den Reichen zusätzliche Steuern aufzubürden. Man mag es für wünschenswert halten, dass das Steuersystem die Konsequenzen ungleicher Lebenschancen korrigiert. Aber soll man die Reichen auch noch dafür bezahlen lassen, dass sie gerne Champagner trinken? Zumal man, wenn man diese Frage bejahen würde, auch dahin käme, höhere Steuern auf Bier zu fordern, sollte sich herausstellen, dass Normalverdiener mehr als alle anderen am Bierkonsum festhalten, wenn sie zu Geringverdienern werden.

Dieses Problem wird intensiv von Benjamin Lockwood und Matthew Weinzierl in einem Aufsatz mit dem sprechenden Titel „De Gustibus non est Taxandum: Heterogeneity in Preferences and Optimal Redistribution“ aus dem Jahr 2014 diskutiert. Sie zeigen auf, dass viele Autoren in ihren Beiträgen zur Theorie optimaler Besteuerung bewusst ein Modell verwendet haben, in dem Haushalte sich nur in ihren produktiven Fähigkeiten und nicht in ihren Konsumpräferenzen unterscheiden. Ein solches Modell mag zwar empirisch falsch sein, vermeidet aber steuerpolitische Schlussfolgerungen, die sich nur dann rechtfertigen lassen, wenn Kon-

11 Die optimale Champagnersteuer wäre nicht-linear, da sie relevant wird, wenn das Einkommen und der Champagnerkonsum ein reduziertes Niveau haben, nicht aber, wenn sie besonders hoch sind, siehe Stiglitz (1982) und Nava et al. (1996).

sumpräferenzen in der Bevölkerung in einer bestimmten Weise verteilt sind, die aber zusammenbrechen, wenn die Verteilung anders aussieht.¹²

Unter der Annahme gleichmäßig verteilter Konsumpräferenzen liefert die Analyse von Atkinson und Stiglitz (1976) eine eindeutige Schlussfolgerung. Verteilungsziele sollten allein mit dem Transfer- und Einkommensteuersystem realisiert werden. Es gibt dann keine Rechtfertigung für differenzierende Verbrauchsteuersätze. Es gibt auch keine Rechtfertigung für einkommensabhängige Kindergartenbeiträge oder für Sozialtarife, die Haushalten helfen, die sonst Schwierigkeiten hätten, ihre Stromrechnung zu bezahlen. Wenn die Einkommensteuer optimal gestaltet ist, dann ist die Verteilungspolitik erledigt. Auch aus ordnungspolitischer Sicht ist das ein attraktives Resultat. Es erlaubt eine klare Zuordnung von steuerpolitischen Instrumenten und steuerpolitischen Zielen.

2.4 Die Besteuerung von Kapitaleinkommen

Wir sind nun einmal durch die klassischen Resultate der Theorie optimaler Steuern gegangen. Ehe wir uns mit den Weiterentwicklungen dieser Theorie in den letzten 10 bis 15 Jahren beschäftigen, lohnt es sich, eine spezielle Anwendung dieser Theorie genauer in den Blick zu nehmen, nämlich ihre Implikationen für die Besteuerung von Kapitaleinkommen.

Dieses Thema ist einerseits durch Thomas Pikettys Buch „Capital in the Twenty-First Century“ im öffentlichen Diskurs sehr präsent. Piketty (2014) dokumentiert, dass die Ungleichheit in der Verteilung der Vermögen überall auf der Welt in den letzten Jahrzehnten drastisch zugenommen hat, und nimmt dies zum Anlass, eine unverteilende Besteuerung von Vermögen zu fordern. Andererseits zeigt sich, dass es ziemlich schwer ist, der klassischen Theorie Argumente für eine positive Besteuerung von Kapitaleinkommen abzurufen. Diese Beobachtung ist der Ausgangspunkt einer jüngeren Literatur, die häufig unter dem Label „New Dynamic Public Finance“ zusammengefasst wird und die wir noch ausgiebiger diskutieren werden.

Die steuerpolitische Praxis unterscheidet zwischen unterschiedlichen Vermögensarten und unterschiedlichen Wegen, zu Vermögen zu gelangen. Grundstücke werden anders behandelt als Anteile an einer Kapitalgesellschaft, die wiederum anders behandelt werden als Ansprüche an eine Rentenversicherung. Auch wird unterschieden zwischen Einkommen aus Zinsen und Dividenden, „legitimeren“ Veräußerungsgewinnen und „weniger legitimen“ Spekulationsgewinnen, sowie zwischen Erbschaften und Schenkungen. Die Theorie der optimalen Besteuerung trifft all diese Unterscheidungen nicht. Sie kennt nur eine Kategorie, die wahlweise als Vermögen, Kapital oder Ersparnis bezeichnet wird. Sie kennt auch nur eine Form der Ein-

12 Die Verwendung eines potentiell falsch spezifizierten Modells, um unplausible Schlussfolgerungen zu vermeiden, ist eine pragmatische Forschungsstrategie. Sie umschifft das Problem, das sich ergibt, weil das zugrundeliegenden Wohlfahrtsmaß unzureichend zwischen Heterogenität aufgrund unterschiedlicher Präferenzen und Heterogenität aufgrund unterschiedlicher produktiver Fähigkeiten unterscheidet. Diese Kritik ist der Ausgangspunkt einer alternativen Formulierung des Problems optimaler Einkommensbesteuerung von Fleurbaey und Maniquet (2006).

kommenserzielung aus diesen Anlagen, die typischerweise als Ertrag aus Kapitalakkumulation oder Ersparnisbildung bezeichnet wird. Diese Form der Abstraktion macht es etwas schwerer, den Zusammenhang zwischen der Theorie und der praktischen Steuerpolitik zu sehen. Sie macht aber auch klar, worum es eigentlich geht: Allen Vermögensarten gemeinsam ist, dass sie über mehrere Perioden Erträge abwerfen und es so erlauben Konsummöglichkeiten aus der Gegenwart in die Zukunft zu verlagern.

Wie würde man auf der Grundlage des Ramsey-Modells über die Besteuerung von Kapitaleinkommen nachdenken? Wir reinterpretieren einen Haushalt, der sein Arbeitsangebot wählt und sich überlegt, wie er sein Einkommen auf Windeln, Champagner und Apfelsinen verteilt, als einen Haushalt, der heute sein Arbeitsangebot wählt und sich überlegt, wie er sein Einkommen auf heutigen Konsum, morgigen Konsum und Konsum in fernerer Zukunft aufteilen will. Zur Erinnerung: Die inverse Elastizitätenregel besagt, dass alle Konsumgüter ohne Ausnahme besteuert werden sollen, dass die Besteuerung aber höher sein sollte, wenn die Nachfrage weniger preissensitiv ist. Demnach sollte also heutiger Konsum besteuert werden aber auch morgiger Konsum und der Konsum im hohen Alter, mit Steuersätzen die an die jeweiligen Elastizitäten angepasst sind. Die Besteuerung künftigen Konsums ist aber nichts anderes als die Besteuerung von Kapitaleinkommen bzw. der Erträge aus Ersparnisbildung.

Diese Intuition legt eine positive Besteuerung von Kapitaleinkommen nahe. Die Arbeiten von Chamley (1986) und Judd (1985) stellen diese Intuition aber in Frage. Beide Aufsätze betrachten ein Ramsey-Modell mit unendlich lange lebenden Haushalten.¹³ Sie zeigen, dass es in einem langfristigen Gleichgewichtszustand keine Besteuerung von Kapitalerträgen geben sollte. Mit folgender Überlegung versucht man typischerweise, dieses Resultats zu plausibilisieren: Eine positive Besteuerung von Kapitalerträgen in einem langfristigen Gleichgewicht hat eine kumulative Wirkung. Von einem Jahr aufs andere schmelzen die Kapitalerträge dahin. Wenn man das ad infinitum fortsetzt, landet man bei Kapitalerträgen von Null. Dies bedeutet auch, dass zukünftiger Konsum unendlich teuer wird. Wenn Haushalte aber eine Präferenz dafür haben, ihren Konsum einigermaßen gleichmäßig auf alle Perioden zu verteilen, dann kann eine Steuerpolitik, die eine langfristige Ersparnisbildung unmöglich macht, nicht optimal sein.

Sowohl die formale Analyse als auch die Interpretation der Resultate von Chamley (1986) und Judd (1985) sind sehr umstritten, und es wäre unangemessen, auf der Grundlage dieser Arbeiten Empfehlungen für die praktische Steuerpolitik abzuleiten, siehe insbesondere die Diskussionen in Atkeson et al. (1999) und Straub und Werning (2015). In jedem Fall liefern diese Arbeiten aber einen ersten Beleg dafür, dass es nicht ganz einfach ist, aus der Theorie optimaler Besteuerung Argumente für eine Besteuerung von Kapitaleinkommen abzuleiten.

13 Man spricht hier auch vom Modell einer dynastischen Familie. Das einzelne Familienmitglied lebt zwar nicht unendlich lange, ist aber in Bezug auf seine Nachkommen derart altruistisch, das sich ein Familienverhalten ergibt, das aussieht, als ob es das Verhalten eines einzelnen unendlich lange lebenden Individuums wäre.

Nun haben wir weiter oben gesehen, dass der Analyserahmen von Ramsey konzeptionelle Schwächen hat. Der Ansatz von Mirrlees (1971) und seine Erweiterung um die Möglichkeit indirekter Besteuerung durch Atkinson und Stiglitz (1976) lassen sich dagegen besser rechtfertigen, jedenfalls sofern man die Annahme trifft, dass die Information über die produktiven Fähigkeiten von Individuen nicht öffentlich verfügbar ist. Was lernen wir also über die optimale Besteuerung von Kapitaleinkommen aus dem Modell von Atkinson und Stiglitz (1976)? Wir reinterpretieren erneut Champagner, Windeln und Apfelsinen als Konsum heute, Konsum morgen und Konsum im hohen Alter. Zur Erinnerung: Atkinson und Stiglitz zeigen, dass eine optimale Steuerpolitik ohne Steuern auf Champagner, Windeln und Apfelsinen auskommt, sofern die Präferenzen für diese Konsumgüter gleichmäßig über alle Einkommensklassen verteilt sind. Damit benötigt man auch keine Besteuerung von Kapitaleinkommen, sofern alle Individuen ein gegebenes Nachsteuereinkommen in ähnlicher Weise auf heutigen Konsum, morgigen Konsum und Konsum in weiter Ferne aufteilen würden. Es wird also nicht leichter, eine Besteuerung von Kapitaleinkommen zu rechtfertigen, wenn man das Ramsey-Modell verlässt und den Analyserahmen von Atkinson und Stiglitz (1976) heranzieht.

3 Jüngere Entwicklungen

Die klassische Literatur lässt zentrale Fragen unbeantwortet. Einige davon hat die jüngere Literatur in den Blick genommen. Hier ist eine Auswahl:

1. Wie hoch sollen optimale Einkommensteuersätze eigentlich sein? Zu wissen, dass sie zwischen Null und 100 Prozent liegen, ist ganz beruhigend, aber für Fragen der praktischen Steuerpolitik nicht allzu hilfreich. Es stellt sich die Frage nach den konkreten quantitativen Implikationen des Modells von Mirrlees.
2. Das Arbeitsmarktmodell, das der Analyse von Mirrlees zugrundeliegt, ist empirisch nicht allzu plausibel. Es unterstellt, dass die Arbeitsintensität eine kontinuierliche Variable ist. Für viele Einkommensbezieher ist die zentrale Frage aber nicht, ob sie 35 oder 45 Stunden pro Woche arbeiten, sondern ob sie überhaupt arbeiten. Außerdem kann eine Subventionierung der Beschäftigungsaufnahme durch Geringqualifizierte im Mirrlees-Modell nicht gerechtfertigt werden. Inwiefern ändert sich das Bild, wenn ein passenderes Arbeitsmarktmodell zugrundegelegt wird?
3. Die klassische Literatur kennt kein Ausland. Sie basiert auf geschlossenen Volkswirtschaften. Es gibt keine Möglichkeit wie Boris Becker Monegasse oder wie Gerard Depardieu Russe zu werden, wenn einem die Steuerpolitik nicht passt. Damit sind wichtige Ausweichreaktionen unzureichend erfasst.
4. Die Steuer- und Transferpolitik des Staates wird in Isolation von der staatlichen Ausgabenpolitik betrachtet. De facto gibt es aber eine Konkurrenzbeziehung zwischen Ausgaben

für Straßen, Schulen oder Polizisten und dem Transfersystem. Wie sieht ein optimales Steuersystem aus, das dieser Interdependenz Rechnung trägt?

5. Die klassischen Analysen berücksichtigen den Lebenszyklus der Steuerzahler nur unzureichend. Individuen kommen nicht fertig ausgebildet auf die Welt und wählen dann eine Arbeitsintensität. Wird die verzerrende Wirkung der Einkommensteuer nicht unterschätzt, wenn Berufs- und Ausbildungsentscheidungen vernachlässigt werden?
6. Individuen haben auch nicht ein und für allemal eine bestimmte Arbeitsproduktivität. Sie sind mit Risiken konfrontiert, können etwa erkranken oder berufsunfähig werden. Das Steuer- und Transfersystem sollte auch eine Versicherung gegen diese Risiken anbieten. Zugleich können Individuen Ersparnisse bilden, um gegen diese Ereignisse besser gewappnet zu sein. Wie interagiert dieses Motiv mit der Besteuerung von Kapitaleinkommen?
7. Die Schlussfolgerung, dass Kapitaleinkommen, Vermögen und Erbschaften nicht besteuert werden sollten, ist für viele Autoren unbefriedigend. Sie widerspricht der steuerpolitischen Praxis fast aller Staaten und suggeriert, dass die Steuerpolitik auf die zunehmende Ungleichheit in der Verteilung der Einkommen und Vermögen nicht reagieren sollte. Dies wirft die Frage nach einer Alternative zu den Analysen von Chamley (1986), Judd (1985) und Atkinson und Stiglitz (1976) auf.
8. Die klassischen Modelle basieren auf der Annahme vollständig rationalen Verhaltens der Steuerzahler. Verhaltensökonomische Studien haben aber zahlreiche Abweichungen von rationalem Verhalten empirisch belegt. Wenn Individuen das Steuersystem nur unvollständig kennen, oder wenn sie Grenz- und Durchschnittssteuersätze nicht auseinanderhalten können, ergeben sich andere Reaktionen auf die Steuerpolitik als vom Rationalmodell vorhergesagt. Welche Implikationen hat dies für die Charakterisierung eines optimalen Einkommensteuertarifs?
9. Die typischerweise verwendeten Wohlfahrtsmaße haben fragwürdige Eigenschaften. Sie basieren auf einem Ideal, wonach produktivere Individuen mehr zum Sozialprodukt beitragen, die Früchte dieser Anstrengung dann aber gleichmäßig an alle verteilt werden. Das einzige, was der vollständigen Verwirklichung dieses vermeintlichen Ideals im Wege steht, sind die Verhaltensreaktionen der Steuerzahler. Das Unbehagen mit dieser Vorstellung führt zur Frage nach alternativen Wohlfahrtsbegriffen und ihren Implikationen für optimale Steuersysteme.
10. Gibt es eine Chance auf die Verwirklichung optimaler Steuerpolitik im politischen Prozess? Wenn die Antwort auf diese Frage „nein“ lautet, welchen praktischen Sinn hat dann die Charakterisierung wohlfahrtsmaximierender Steuersysteme? Um diese Frage zu klären ist es erforderlich, sich mit der politischen Ökonomie umverteilender Steuern zu beschäftigen.

Wir werden Antworten auf diese Fragen besprechen. Allerdings haben nicht alle diese Fragen den gleichen Raum in der jüngeren Literatur eingenommen. In Abschnitt 3.1 besprechen wir die jüngeren Entwicklungen zur Quantifizierung optimaler Steuern, d.h. die Antwort auf Frage 1. Die Abschnitte 3.2-3.4 widmen sich den Fragen 2 bis 4. Diesen Beiträgen ist gemeinsam, dass sie die Analyse von Mirrlees um weitere Formen individueller Heterogenität erweitern. Es gibt also nicht mehr nur Unterschiede in den produktiven Fähigkeiten sondern beispielsweise auch in der Zahl zu versorgender Kinder oder der Bereitschaft ins Ausland zu ziehen. Es wird auch eine größere Zahl von Entscheidungen analysiert. Neben die Betrachtung des Arbeitseinsatzes tritt etwa die Wahl des Wohnsitzes oder die Nachfrage nach öffentlichen Gütern. Die Fragen 5, 6 und 7 betreffen die Wirkung der Besteuerung im Lebenszyklus. Ausbildungsentscheidung und Berufswahl stehen am Anfang und die Vererbung von Vermögen am Ende des Lebens. Die Erweiterung des Mirrlees-Modells um das Thema der Absicherung gegen den Verlust produktiver Fähigkeiten (Frage 7) hat eine Literatur beschäftigt, die unter der Überschrift „New Dynamic Public Finance“ firmiert. Diese Entwicklungen werden in Abschnitt 3.5 aufgegriffen.

Die Fragen 8-10 betreffen die Grundlagen der Theorie optimaler Besteuerung. Auch hierzu sind in den vergangenen Jahren einige wichtige Arbeiten entstanden. Diese Literatur ist meines Erachtens aber noch in einem explorativen Stadium, so dass ich noch keine Möglichkeit einer griffigen Systematisierung sehe. Es ist aber absehbar, dass diese Themen in den nächsten Jahren weiter vertieft werden. Ich greife sie daher in den abschließenden Bemerkungen erneut auf.

3.1 Quantifizierung optimaler Steuern

Mit den Arbeiten von Diamond (1998) und Saez (2001) wird die Literatur zur optimalen Einkommensbesteuerung sehr viel konkreter. An die Stelle theoretischer Forschung zum Vorzeichen optimaler Grenzsteuersätze tritt vermehrt die Quantifizierung. Diamonds Aufsatz enthält zwei zentrale Einsichten, eine theoretische und eine empirische. Die theoretische Einsicht ist als „ABC-Formel“ bekannt geworden: Wenn man von Einkommenseffekten im Konsum absieht, lassen sich optimale Einkommenssteuersätze einfach darstellen. Der Grenzsteuersatz bei einem Einkommen von y sollte umso höher sein, je

- A: günstiger das Verhältnis von Haushalten, die mehr Steuern zahlen, zu Haushalten ist, die weniger arbeiten wenn dieser Steuersatz erhöht wird,
- B: geringer die Ausweichreaktion derer ist, die weniger arbeiten, und
- C: weniger das verwendete Wohlfahrtsmaß die Interessen derer berücksichtigt, die ein Einkommen oberhalb von y erzielen.

Dabei kann B gemessen werden durch die Elastizität des Arbeitsangebots in Bezug auf den Nettolohn, und A gemessen werden durch eine spezielle Eigenschaft der Einkommensvertei-

lung, die sogenannte Hazard-Rate. Zur Erläuterung der Hazard Rate rufen wir uns erneut die oben angestellte Überlegung zu einem höheren Steuersatz für alle Einkommen zwischen 75.000 und 76.000 Euro in Erinnerung. Alle Individuen mit einem Einkommen zwischen 75.000 und 76.000 Euro haben dann schlechtere Anreize und alle Individuen, die mehr als 75.000 Euro verdienen zahlen mehr Steuern. Sei f der Bevölkerungsanteil der Individuen mit einem Einkommen zwischen 75.000 und 76.000 Euro, und sei F der Bevölkerungsanteil aller Individuen mit einem Einkommen unter 75.000 Euro, und folglich $1-F$ der Anteil aller Individuen mit einem Einkommen jenseits der 75.000. Die Hazard Rate ist definiert als $1-F/f$, also als das Verhältnis derer, die mehr Steuern zahlen, zu jenen, die schlechtere Anreize haben. Je größer dieser Ausdruck ist, desto attraktiver ist ein hoher Steuersatz.

Die empirische Einsicht ist, dass zumindest im Bereich hoher Einkommen die Einkommensverteilung gut durch eine Pareto-Verteilung approximiert werden kann. Die Pareto-Verteilung hat insbesondere die Eigenschaft, dass die Hazard-Rate konstant ist. Wenn man sich zunächst von den niedrigen zu den mittleren Einkommen bewegt und dann von den mittleren zu den hohen und schließlich von den hohen zu den noch höheren Einkommen, immer bleibt der Quotient $1-F/f$ der gleiche. Diamond zeigt, dass das auch für die empirische Einkommensverteilung gilt, jedenfalls solange man nur hohe Einkommen in den Blick nimmt.

Zur Berechnung eines optimalen Einkommenssteuertarifs unterstellt Diamond, dass die Elastizität des Arbeitsangebots eine einkommensunabhängige Konstante darstellt (eine Annahme, die in der angewandten Forschung aus pragmatischen Gründen häufig getroffen wird) und dass die hohen Einkommen einer Pareto-Verteilung folgen. Er geht noch etwas weiter und unterstellt, dass die Pareto-Verteilung auch für die Verteilung der Spitzeneinkommen noch eine gute Approximation darstellt.¹⁴ Dann braucht er nur noch eine Spezifikation des Wohlfahrtsmaßes und kann unter Verwendung der ABC-Formel optimale Grenzsteuersätze ausrechnen. Insbesondere ergibt sich so eine Rechtfertigung steigender Grenzsteuersätze auch im Bereich der hohen Einkommen: Wenn A und B tatsächlich konstant bleiben, wenn wir uns entlang der Einkommensverteilung von den hohen zu den sehr hohen Einkommen bewegen, und das Wohlfahrtsgewicht der „Reichen“ höher ist als das der „sehr Reichen“, dann folgt aus der ABC-Formel, dass der Grenzsteuersatz für die „sehr Reichen“ höher sein sollte als für „die Reichen“. Wenn man dann irgendwann in einen Bereich gelangt, wo die Wohlfahrtsgewichte der Reichen nicht weiter sinken, konvergieren auch die Grenzsteuersätze. Unter dieser Annahme rechnet Diamond optimale Grenzsteuersätze im Bereich der sehr hohen Einkommen explizit aus und zeigt, dass sie zwischen 60 und 90 Prozent liegen, also weit über den Spitzensteuersätzen, die man in der Realität beobachtet. Diamonds Analyse liefert also auch quantitativ ein eindeutiges Resultat: Aus der Perspektive seines Modells, werden die Grenzsteuersätze für die Bezieher hoher Einkommen in der steuerpolitischen Praxis zu niedrig gesetzt.

14 Die Ergebnisse der Analyse hängen stark von der Annahme einer Pareto-Verteilung ab. Unterstellt man dagegen, dass die Einkommen lognormal verteilt sind, ergibt sich ein völlig anderes Bild, siehe Mankiw et al. (2009).

Wie passt das zusammen mit dem weiter oben geschilderten Resultat, wonach Spitzenverdiener einen Grenzsteuersatz von Null haben sollten? Diamond vermeidet diese Schlussfolgerung durch die Annahme, dass die Einkommensverteilung durch eine Pareto-Verteilung gut approximiert werden kann. Die Pareto-Verteilung kennt aber keinen Spitzenverdiener, es ist eine nach oben nicht beschränkte Verteilung. Daher gibt es zu jedem Einkommensbezieher jemanden, der noch mehr Einkommen hat. Anders ausgedrückt: Es gibt keinen Spitzenverdiener. De facto muss aber in einem Land mit endlich vielen Einwohnern irgendwer der Reichste sein. Der Bereich der Einkommensverteilung, der unmittelbar auf den Spitzenverdiener zuläuft, lässt sich daher nicht mehr durch eine Pareto-Verteilung approximieren. In diesem Bereich ist Diamonds Ansatz also nicht angemessen. Stattdessen gilt das Resultat, wonach der Spitzenverdiener einen Grenzsteuersatz von Null haben sollte.¹⁵

Die Arbeit von Diamond (1998) ist zum Ausgangspunkt weiterer Analysen geworden. In Saez (2001) werden zusätzlich Einkommenseffekte berücksichtigt. Auch die Darstellung optimaler Einkommensteuertarife gemäß der „ABC-Formel“ hat zahlreiche Erweiterungen erfahren. Beispielsweise betrachten Kleven et al. (2009) die optimale Besteuerung von Ein- und Mehrverdienerhaushalten.¹⁶ Die weiter unten diskutierten Arbeiten von Rothschild und Scheuer (2013), die die Berufswahl berücksichtigen, von Lehmann et al. (2014), die auf die Möglichkeit des Wegzugs von Steuerzahlern ins Ausland eingehen oder von Jacquet et al. (2013), die ein alternatives Arbeitsmarktmodell zugrunde legen, nehmen alle Diamonds ABC-Formel zum Ausgangspunkt.

3.2 Arbeitsmarktpartizipation

Das Arbeitsmarktmodell, auf dem die Analyse von Mirrlees basiert, unterstellt, dass Haushalte ihr Arbeitsangebot in Reaktion auf Änderungen von Verdienstmöglichkeiten oder Steuersätzen variabel anpassen. Auf dieser Annahme basiert das Resultat, wonach eine optimale Steuerpolitik das Arbeitsangebot der Steuerzahler nach unten verzerrt. Die Steuerzahler arbeiten also weniger als in einer hypothetischen Laissez-Faire-Ökonomie ohne Besteuerung.

Die Erfahrung persistenter Arbeitslosigkeit in vielen OECD-Staaten wirft die Frage auf, ob Verzerrungen nach unten wirklich wünschenswert sind. Sozialpolitische Programme, wie etwa der Earned-income-tax-credit in den USA, die auf einer Subventionierung der Beschäftigungsaufnahme von Geringverdienern basieren, entsprechen einem System mit negativen Grenzsteuersätzen, da die Beschäftigten zusätzlich zu ihrem Stundenlohn eine Transferzahlung erhalten. Diese Programme sind nicht mit einem optimalen Steuersystem à la Mirrlees vereinbar.

Zugleich belegt die empirische Arbeitsmarktforschung, dass die Verhaltensreaktionen auf veränderte Arbeitsanreize durch das Standardmodell unzureichend erfasst werden, siehe etwa

15 Für eine explizitere Betrachtung der Spitzenverdiener siehe Scheuer und Werning (2015).

16 Siehe auch Bach et al. (2012) für eine Anwendung mit deutschen Daten.

Blundell et al. (2011). Viele Haushalte passen ihr Arbeitsangebot nicht kontinuierlich an, sondern springen zwischen Beschäftigungslosigkeit und Beschäftigung hin und her. Diese Verhaltensreaktion lässt sich durch ein Modell, in dem Berufstätigkeit mit fixen Kosten verbunden ist, besser abbilden. Um ein Beispiel zu geben: Wenn das Einkommen ausreicht, um die im Falle der Berufstätigkeit erforderliche Kinderbetreuung zu finanzieren, springt das Arbeitsangebot vieler Arbeitnehmer von Null auf 30 Stunden pro Woche. Das Mirrlees-Modell kennt dagegen nur graduelle Verhaltensanpassungen, etwa einen Anstieg der Arbeitszeit von 14 auf 15 Stunden.

Beginnend mit einer Arbeit von Saez (2002) haben sich viele Aufsätze mit der Frage beschäftigt, wie die Berücksichtigung solcher diskontinuierlicher Verhaltensreaktionen die Eigenschaften eines optimalen Steuer- und Transfersystems beeinflusst und insbesondere, ob sich auf dieser Grundlage auch eine Politik mit Beschäftigungsanreizen für Geringverdiener rechtfertigen lässt. Die Literatur hat zwei Argumente identifiziert, die für eine solche Politik sprechen.

Das erste Argument stellt auf die Intensität der Verhaltensreaktion ab, siehe Jacquet et al. (2013) und Hansen (2015). Diese fällt in einem Modell mit fixen Kosten aus Erwerbstätigkeit drastischer aus als im Mirrlees-Rahmen. Von Mirrlees wissen wir, dass positive Steuersätze hilfreich sind, um Einkommen von gut verdienenden Haushalten zu schlechter verdienenden Haushalten umzuverteilen. In einem Modell mit fixen Kosten der Arbeitsmarktpartizipation drängen solche Steuern zusätzlich Menschen aus dem Arbeitsmarkt. Dies steigert die Wohlfahrtskosten der Besteuerung. Damit wird das Argument für positive Steuersätze schwächer. Wirkt der Partizipationseffekt sehr stark, werden sogar negative Steuersätze wünschenswert.

Das zweite Argument basiert auf der Beobachtung, dass Steuerzahler sich nicht nur in ihren produktiven Fähigkeiten unterscheiden, sondern auch in einer Reihe anderer Charakteristika, etwa der Zahl der Kinder, oder ihrem Gesundheitszustand. Dies kann dazu führen, dass die Bedürftigkeit eines Haushaltes nicht mehr perfekt durch die Höhe des Vorsteuereinkommens gemessen werden kann. Betrachten wir als Beispiel den Vergleich zwischen Transferempfängern ohne eigenem Einkommen („non-working poor“) und Geringverdienern („working poor“). Wenn wir unterstellen, dass diese sich nur in ihren Fähigkeiten der Einkommenserzielung unterscheiden, dann schließen wir, dass es den „non-working poor“ noch einmal schlechter gehen muss als den „working-poor.“ Diese Sichtweise ist passend, wenn wir uns die „non-working poor“ als Menschen mit körperlichen oder geistigen Beeinträchtigungen vorstellen und die „working-poor“ als gesunde, aber ungelernete Arbeitskräfte. Wenn wir uns dagegen vorstellen, dass die „working poor“ sich von den „non-working poor“ dadurch unterscheiden, dass der Druck der Einkommenserzielung für sie besonders hoch ist, etwa weil sie Kinder zu versorgen haben, dann ist die Bedürftigkeit der „working poor“ höher als die der „non-working poor.“ Die bedürftigere Gruppe ist dann also die mit dem höheren Vorsteuereinkommen. Die Arbeiten von Saez (2002) sowie von Chone und Laroque (2011) haben herausgearbeitet, dass eine besondere Bedürftigkeit der „working poor“ für eine Subventionierung der Geringverdiener spricht. Im typischen Analyserahmen von Mirrlees ist dies nicht denkbar:

Wenn die Bedürftigkeit monoton im Einkommen fällt, dann sind durchweg positive Grenzsteuersätze optimal. Wenn aber im Übergang von den „non-working poor“ zu den „working poor“ die Bedürftigkeit steigt, dann sind negative Grenzsteuersätze optimal.

Wir haben weiter oben eine Unterscheidung getroffen zwischen pragmatischen Forschungsansätzen und Forschungsansätzen, die auf einer vollständigen allokatorentheoretischen Begründung bestehen. Wohin gehört die Literatur zur Subventionierung der Beschäftigungsaufnahme? Diese Frage ist nicht ganz einfach zu beantworten: Typischerweise beginnen diese Arbeiten mit der Feststellung, dass der klassische Analyserahmen von Mirrlees unzureichend ist, weil er nur eine Form von Heterogenität kennt; Individuen unterscheiden sich nur hinsichtlich der Frage, wie viel Aufwand sie treiben müssen, um etwas mehr Einkommen zu erzielen. Dann wird argumentiert, dass Individuen sich aber auch anhand anderer Kriterien unterscheiden, die wichtig für die Entscheidung sind, überhaupt eine reguläre Beschäftigung anzustreben. Als Beispiele werden genannt: die Zahl der Kinder, die Frage, ob man Angehörige pflegen muss, welche Ausbildung man hat, wie gesund man ist, etc. Auf dieser Grundlage wird dann ein theoretischer Analyserahmen präsentiert, der unterstellt, dass all die Ausprägung dieser Heterogenität private Information darstellt und ein optimales Steuer- und Transfersystem hergeleitet. Anders gesagt: Wenn es wirklich der Fall wäre, dass der Staat keine Information darüber hat, wie gesund die Steuerzahler sind, wie viele Kinder sie haben, über welchen Schulabschluss sie verfügen oder ob sie Angehörige pflegen, dann gäbe es eine allokatorentheoretische Begründung für die Subventionierung der Beschäftigungsaufnahme. Diese Annahme macht aber keinen Sinn, wenn es ein öffentliches Schulsystem gibt, ein öffentliches Gesundheitswesen, eine gesetzliche Pflegeversicherung und ein Steuersystem, bei dem die privaten Ausgaben für Bildung, Gesundheit und Pflege bei der alljährlichen Steuererklärung angegeben werden. Wenn diese Informationen aber öffentlich verfügbar sind, dann wird ein optimales Steuer- und Transfersystem sie auch nutzen. In der Konsequenz hieße das: Es gäbe einen eigenen Einkommensteuertarif für Menschen mit Kindern und für Menschen ohne Kinder, es gäbe einen eigenen Tarif für Menschen, die Angehörige pflegen, und für Menschen, die das nicht tun.¹⁷ Man bildet also Teilgruppen der Bevölkerung. Innerhalb jeder Teilgruppe gibt es dann nur noch eindimensionale Heterogenität, nämlich die im Mirrlees-Modell unterstellte Heterogenität in den produktiven Fähigkeiten. In der Konsequenz gibt es auch keine Rechtfertigung mehr für Subventionen, die auf die Beschäftigungsaufnahme zielen.

Diese Argumentation zeigt, dass die Subventionierung von Geringverdienern nur dann wünschenswert ist, wenn man zugleich unterstellt, dass eine Steuerpolitik, die für die oben genannten Teilgruppen einen jeweils eigenen Einkommensteuertarif vorsieht, nicht zur Diskussion steht. Die Empfehlung, Beschäftigung zu subventionieren kann also nur durch einen

17 Im Jargon der Literatur spricht man von „Tagging“ wenn Steuern und Transferzahlungen auf solche beobachtbaren Charakteristika der Steuerzahler konditionieren. Diese Terminologie geht zurück auf einen Aufsatz von Akerlof (1978), der gezeigt hat, dass man "Tagging" betreiben sollte wann immer möglich, schlicht weil es beim Konditionieren auf beobachtbare Charakteristika keine Ausweichreaktionen der Steuerzahler gibt.

pragmatischen Forschungsansatz gestützt werden kann, der auf der Annahme basiert, dass das „wahre Optimum“ aus Gründen, die unerklärt bleiben, ohnehin nicht zur Debatte steht.

3.3 Mobile Steuerzahler

Wie ändert sich die Betrachtung optimaler Steuersysteme, wenn wir eine weitere Verhaltensreaktion ins Kalkül einbeziehen, nämlich die Möglichkeit in ein anderes Land zu ziehen? Wir können zunächst die optimale Steuerpolitik eines Landes betrachten, das diese Ausweichreaktionen gebührend berücksichtigt. Wir können aber darüber hinaus auch das Verhalten mehrerer Regierungen gleichzeitig betrachten, die versuchen, einander begehrte Steuerzahler abzuluchsen und fragen, welche Ergebnisse unter den Bedingungen eines solchen Steuerwettbewerbs zu erwarten sind. Eine gängige Hypothese ist, dass der Steuerwettbewerb ein „Race-to-the-bottom“ zur Folge hat, so dass am Ende die Steuer- und Transfersysteme erodieren. Regierungen geraten demnach in einen Unterbietungswetlauf und müssen, um überhaupt Steuerzahlungen generieren zu können, den leistungsfähigen Steuerzahlern der globalisierten Weltwirtschaft immer niedrigere Steuersätze anbieten.

Das Phänomen des Steuerwettbewerbs ist zuerst in einem anderen Kontext studiert worden, nämlich als Konkurrenz verschiedener Standorte um die Ansiedlung von Firmen. Die öffentlichen Auseinandersetzungen um die niedrigen Steuern, die Google, Starbucks oder Apple zahlen, oder die Verlagerung von Unternehmenssitzen in Steueroasen wie die Bahamas oder die Cayman Inseln, liefern hierzu das Anschauungsmaterial. Inwiefern ist Steuerwettbewerb auch ein Thema für die Analyse von Einkommensteuern? Dies ist ein jüngeres Thema der Literatur, aber es gibt zumindest einige Hinweise, dass sich diese Form von Steuerwettbewerb auch empirisch nachweisen lässt.¹⁸ Die Frage nach den möglichen Konsequenzen des Einkommensteuerwettbewerbs ist auch interessant im Hinblick auf die immer wieder geführte Diskussion, ob die deutschen Bundesländer eigene Gestaltungsmöglichkeiten bei der Einkommensbesteuerung haben sollten, siehe etwa die Diskussion in Schulte (2014). Die Arbeiten von Bierbrauer et al. (2013) und Lehmann et al. (2014) gehen dieser Frage auf der Grundlage von theoretischen Überlegungen nach. Lehmann et al. (2014) unterstellen, dass Haushalte sich in ihren produktiven Fähigkeiten unterscheiden wie im konventionellen Mirrlees-Modell und in den Kosten, die der Umzug in ein anderes Land für sie verursachen würde.¹⁹ Es wird eine Situation betrachtet, in der die Regierungen mehrerer Länder eine umverteilende Steuerpolitik wählen und wechselseitig beste Antworten finden müssen, im Gegensatz zur Situation in der ein Land in Isolation vom Rest der Welt agiert. Die Autoren zeigen, dass in einem spieltheoretischen Gleichgewicht die Steuersätze und das Transfervolumen niedriger sind als im Modell einer geschlossenen Volkswirtschaft. Allerdings gibt es kein vollständiges

18 Kleven et al. (2014) beschäftigen sich mit den Auswirkungen spezieller Steuervorteile, die Ausländer in Dänemark genießen, und Kleven et al. (2013) analysieren die Auswirkungen unterschiedlicher nationaler Steuerpolitiken auf die Mobilität von Fußballprofs.

19 Der Analyserahmen ist daher ähnlich wie in dem weiter oben erwähnten Aufsatz von Jacquet et al. (2013), in dem Haushalte entscheiden, wie intensiv sie arbeiten, und fixe Kosten durch Erwerbstätigkeit entstehen. Hier sind die fixen Kosten mit dem Wechsel des Wohnsitzes verbunden.

„Race-to-the-bottom“, da die Haushalte nicht vollständig mobil sind, so dass die Wohlfahrtskosten der Umverteilung relativ zum Modell einer geschlossenen Volkswirtschaft zwar steigen, aber nicht so hoch werden, dass es überhaupt nicht mehr zu umverteilender Steuerpolitik kommt. Bierbrauer et al. (2013) betrachten dagegen den theoretischen Extremfall vollständig mobiler Haushalte, die friktionslos an den Standort ziehen, der ihnen die attraktivere Steuerpolitik bietet. Dies verändert das spieltheoretische Gleichgewicht dahingehend, dass umverteilende Steuerpolitik unmöglich wird. Stattdessen kann es sogar dazu kommen, dass Regierungen von unten nach oben umverteilen, um einkommensstarke Haushalte in ihr Land zu locken. Der Anreiz dazu besteht, weil die nationale Wohlfahrt steigt, wenn mehr einkommensstarke Haushalte im Land sind.

Was kann man aus diesen Aufsätzen lernen? Um diese Frage zu beantworten, rufen wir uns zunächst andere Vorstellungen zu den Wirkungen des Wettbewerbs zwischen verschiedenen Jurisdiktionen in Erinnerung. Mit Arbeiten in der Tradition von Tiebout (1956) ist die Vorstellung verbunden, dass dieser Wettbewerb zu einer effizienten Aufteilung von Bürgern auf verschiedene Standorte führt, siehe auch Jehiel und Lamy (2015) für eine jüngere Formalisierung dieser Hypothese. In Tiebouts Vorstellung gibt es Bürger mit einer hohen Zahlungsbereitschaft für öffentliche Güter (Straßen, Schulen, Theater) und solche mit einer niedrigen Zahlungsbereitschaft. Wenn die einen an einen Ort mit niedrigen Steuern und einer geringen Bereitstellung öffentlicher Güter ziehen und die anderen an einen Ort mit höheren Steuern und mehr öffentlichen Gütern, entsteht eine effiziente Situation. Die Public-Choice Schule, siehe zum Beispiel Brennan und Buchanan (1980), betrachtet den öffentlichen Sektor als einen gefräßigen Leviathan, der sein Gewaltmonopol ausnutzt, um die Interessen von Bürokraten, Politikern und Lobbyisten zu bedienen und nicht in erster Linie Gemeinwohlziele verfolgt. Aus Sicht dieser Theorie ist es ein wünschenswerter Effekt, wenn Steuerwettbewerb dazu führt, dass die Basis der Besteuerung erodiert. Unterschiedliche Präferenzen der Bürger für öffentliche Güter oder Agentur-Probleme zwischen politischen Akteuren und Bürgern lassen den Steuerwettbewerb also in einem günstigen Licht erscheinen. Vor diesem Hintergrund zeigen die Arbeiten von Bierbrauer et al. (2013) und Lehmann et al. (2014), dass es aber eine Dimension staatlichen Handelns gibt, die durch den Steuerwettbewerb gefährdet ist, nämlich die Absicherung von Bürgern gegen Armut und elementare Lebensrisiken. Eine umverteilende Steuerpolitik wird durch den Steuerwettbewerb erschwert.

3.4 Optimale Steuern und die optimale Bereitstellung öffentlicher Güter

Die Theorie der optimalen Besteuerung behandelt typischerweise die Staatsausgaben als ein Datum. Die Frage nach den optimalen Staatsausgaben und die Frage nach der optimalen Gestaltung der Staatseinnahmen werden also getrennt voneinander behandelt. Die Analyse der Ausgabenseite bestimmt dabei, was an Staatseinnahmen erforderlich ist, um die Ausgaben finanzieren zu können. Dieses Einnahmeziel findet dann als Nebenbedingung Eingang in die Analyse optimaler Steuern.

Dieses Vorgehen ist aus zwei Gründen problematisch. Erstens hängt das optimale Bereitstellungsniveau öffentlicher Güter von den Wohlfahrtskosten der Besteuerung ab. Die Wohlfahrtskosten der Besteuerung hängen wiederum davon ab, welche Instrumente der Besteuerung zur Verfügung stehen: Es macht einen Unterschied, ob man in der Analyse verhaltensunabhängige Pauschalsteuern, nicht-lineare Einkommensteuern à la Mirrlees (1971) oder ein lineares Verbrauchsteuersystem à la Ramsey (1927) zugrunde legt. Die Frage nach den optimalen Staatsausgaben lässt sich also nicht unabhängig vom Steuersystem beantworten. Zweitens hängen auch die individuellen Präferenzen für öffentliche Güter vom Steuersystem ab. Steuerzahler, die fürchten müssen, stark zur Finanzierung zusätzlicher Ausgaben herangezogen zu werden, denken über einen Ausbau des Straßennetzes anders nach als Individuen, die dafür nicht in Frage kommen. Dies schafft Anreize zu strategischem Verhalten im politischen Prozess mit dem Ziel, die Ausgabenentscheidungen im jeweils eigenen Interesse zu beeinflussen, und wirft die Frage auf, wie sicher gestellt werden kann, dass bei der Kosten-Nutzen-Analyse zur Bereitstellung öffentlicher Güter auf ein unverzerrtes Maß der Zahlungsbereitschaft für öffentliche Güter zurückgegriffen werden kann.

Die Wohlfahrtskosten der Besteuerung und die Bereitstellung öffentlicher Güter. Das zuletzt genannte Problem entsteht nicht, wenn die Präferenzen für öffentliche Güter als bekannt unterstellt werden. Unter dieser vereinfachenden Annahme kann sich die Analyse auf die Frage konzentrieren, wie die Wohlfahrtskosten der Besteuerung bei der Gestaltung der staatlichen Ausgabenpolitik berücksichtigt werden sollten. Gaube (2000) unterstellt, dass ein lineares System von Verbrauchsteuern, in der Formalisierung von Diamond (1975), zur Finanzierung öffentlicher Güter herangezogen wird und stellt die Frage, ob dies dazu führt, dass – relativ zu einer Situation, in der die Finanzierung über unverzerrende Pauschalsteuern erfolgt – weniger öffentliche Güter bereitgestellt werden. Die gängige Intuition ist, dass höhere Wohlfahrtskosten der Besteuerung auch weniger Staatsausgaben nach sich ziehen sollten. Gaube zeigt anhand mehrerer Gegenbeispiele, dass diese Vorstellung falsch ist: Aus der Verwendung verzerrender Steuern zur Staatsfinanzierung lässt sich nicht logisch folgern, dass der Umfang staatlicher Aktivitäten kleiner werden sollte.

Boadway und Keen (1993) betrachten ein Modell, in dem ein nicht-lineares Einkommensteuersystem à la Mirrlees (1971) verwendet wird, um sowohl Transfers zu Geringverdienern als auch die Bereitstellung öffentlicher Güter zu finanzieren. Die zentrale Fragestellung ist, ob die Bereitstellung öffentlicher Güter weiter dem Effizienzkriterium folgen sollte, das als Samuelson-Regel bekannt geworden ist (Samuelson 1954). Nach diesem Effizienzkriterium soll die Bereitstellung öffentlicher Güter solange ausgedehnt werden, bis die kollektive Zahlungsbereitschaft – definiert als die Summe der Zahlungsbereitschaften aller, die von der Bereitstellung öffentlicher Güter profitieren – gerade dem marginalen Ressourcen aufwand entspricht. Nach dieser Regel ergibt sich das relevante Kostenmaß also gerade nicht aus den Wohlfahrtskosten der Besteuerung, sondern nur aus den technologischen Gegebenheiten der Ökonomie. Boadway und Keen (1993) zeigen, dass unter bestimmten Annahmen an die Präferenzen diese Effizienzregel weiter gelten sollte, obwohl verzerrende Steuern zur Finanzierung öffentlicher

Ausgaben verwendet werden. Die Logik ist der von Atkinson und Stiglitz (1976) verwandt. Atkinson und Stiglitz zeigen, dass es keine Rolle für verzerrende Verbrauchsteuern gibt, wenn ein optimales Einkommensteuersystem zur Staatsfinanzierung herangezogen wird. Analog gibt es unter diesen Bedingungen auch keine Rolle für eine Verzerrung des Angebots an öffentlichen Gütern.

Halten wir also fest: Weder die Wohlfahrtskosten indirekter Verbrauchsteuern noch diejenigen direkter Einkommensbesteuerung legen a priori nahe, dass, relativ zu einer Situation mit unverzerrenden Pauschalsteuern, eine Unterversorgung mit öffentlichen Gütern angezeigt wäre.

Die Analyse von Boadway und Keen (1993) lässt allerdings folgende Frage unbeantwortet: Wie interagieren Verteilungsziele mit der Bereitstellung öffentlicher Güter? Anders formuliert, wenn wir von einem Wohlfahrtsmaß übergehen zu einem anderen, das ein stärkeres Gewicht auf das Wohlergehen der Geringverdiener legt, wie verändert sich dann das optimale Bereitstellungs-niveau öffentlicher Güter? Eine Antwort auf diese Frage findet sich in Bierbrauer und Boyer (2014). Eine stärkere Betonung des Verteilungsziels reduziert das optimale Bereitstellungs-niveau. Geringverdiener leiden stärker, wenn mehr Ressourcen für öffentliche Güter benötigt werden und nicht mehr für Umverteilung zur Verfügung stehen. Dieser Effekt wird stärker berücksichtigt, wenn Verteilungsziele mehr Gewicht erhalten.

Heterogene Präferenzen für öffentliche Güter. Die Arbeiten von Boadway und Keen (1993) und Bierbrauer und Boyer (2014) basieren auf der Annahme, dass Individuen sich nur in ihren produktiven Fähigkeiten unterscheiden, nicht jedoch in ihren Präferenzen für öffentliche Güter. In diesen Modellen ist der Fall ausgeschlossen, dass zwei Individuen das gleiche Einkommen erzielen aber unterschiedliche öffentliche Güter, etwa das Straßennetz und das Schienennetz, nutzen wollen. Hellwig (2004) erweitert diese Analyse und lässt Heterogenität sowohl in den produktiven Fähigkeiten als auch in den Präferenzen für öffentliche Güter zu. Er betrachtet ausschließbare öffentliche Güter, also solche, deren Nutzung an das Zahlen eines „Eintrittspreises“ geknüpft werden kann, etwa an die Entrichtung einer Maut oder den Erwerb einer Fahrkarte. Damit entsteht ein Analyserahmen, in dem ein nicht-lineares Einkommensteuersystem und ein System linearer Eintrittspreise für öffentliche Güter zur Verfügung stehen, um sowohl das Transfersystem als auch die Bereitstellung öffentlicher Güter zu finanzieren.²⁰ Hellwig zeigt, dass ausgeprägte Verteilungsziele es wünschenswert machen, von einer Bereitstellung öffentlicher Güter gemäß der Samuelson-Regel abzuweichen. Stattdessen soll es zu einer Verknappung des Angebots an öffentlichen Gütern kommen, was es möglich macht, höhere Nutzungsentgelte zu verlangen. Die Einnahmen aus den Nutzungsent-

20 Die Analyse kombiniert daher ein Problem der Preissetzung im öffentlichen Sektor in der Tradition von Boiteux (1956) mit der Betrachtung nicht-linearer Einkommensteuern. Ein ähnliches Preissetzungsproblem ergibt sich auch, wenn Einkommensteuern nicht zur Verfügung stehen, aber den öffentlichen Unternehmen die Pflicht auferlegt wird, sich selbst über Nutzungsentgelte im Gegensatz zur Kostendeckung aus dem Staatshaushalt zu refinanzieren siehe (Hellwig 2007b). Inwiefern es wünschenswert ist, öffentlichen Unternehmen eine solche Selbstfinanzierungspflicht aufzuerlegen, wird in Bierbrauer (2011b) diskutiert.

gelten können ins Transfersystem gespeist werden, und dies wiederum erleichtert das Erreichen von Verteilungszielen.

Ein kritischer Punkt in Hellwigs Analyse ist, dass die Heterogenität in den Präferenzen für öffentliche Güter eine weitere Dimension eröffnet, in der es eine Rolle für Verteilungspolitik gibt. Es geht nicht mehr nur um Umverteilung zwischen „arm“ und „reich“, sondern auch um Umverteilung zwischen denjenigen, die stark vom öffentlichen Straßennetz profitieren, zu jenen, die das nicht tun.²¹ Ein wohlfahrtsmaximierender Ansatz macht es wünschenswert, die derart Privilegierten zur Kasse zu bitten, um ihren Vorteil gegenüber jenen, die weniger stark profitieren, möglichst gering zu halten. Ohne diese Form von Präferenzheterogenität wären wir im Analyserahmen von Boadway und Keen (1993), mit der Konsequenz, dass eine Abweichung vom Effizienzmaßstab der SamuelsonRule nicht angezeigt wäre.

Die Vorstellung, dass diejenigen besonders zur Finanzierung der Staatsaufgaben herangezogen werden sollten, die am stärksten profitieren, wird häufig auch als „Benefits-based Taxation“ bezeichnet und geht auf Lindahl (1919) zurück. Für Lindahl war sie ein eigenständiges Prinzip in der normativen Theorie der Besteuerung. Hellwigs Analyse zeigt, dass eine Besteuerung in Abhängigkeit davon, wie sehr Individuen von der Bereitstellung öffentlicher Güter profitieren, aus einem konventionellen Wohlfahrtskalkül folgt. Es ist demnach nicht erforderlich, „Benefits-based Taxation“ als ein eigenständiges Prinzip gerechter Besteuerung zu rechtfertigen. Eine jüngere Arbeit von Weinzierl (2014b) dokumentiert jedoch anhand von zahlreichen Äußerungen im öffentlichen Diskurs, dass viele Menschen dies als ein eigenständiges normatives Prinzip zu empfinden scheinen.

Die Wohlfahrtskosten der Besteuerung und die Aggregation von Präferenzen für öffentliche Güter. Die Arbeiten von Bierbrauer (2009, 2012, 2014) unterstellen, dass die Nachfrage nach öffentlichen Gütern nicht a priori bekannt ist. In der Konsequenz muss der Entscheidung über die Bereitstellung öffentlicher Güter eine Aggregation individueller Präferenzen vorausgehen. Dies erzeugt eine Möglichkeit strategischer Kommunikation, um die Ergebnisse des politischen Prozesses zu beeinflussen. Eine zentrale Beobachtung in den genannten Arbeiten ist, dass es vom Steuersystem selbst abhängt, in welche Richtung die Anreize für eine solche strategische Kommunikation gehen. Wird ein einheitliches Verbrauchsteuersystem zur Finanzierung öffentlicher Güter verwendet, werden gutverdienende Haushalte die Hauptlast der Finanzierung tragen. Im Prozess der politischen Willensbildung haben Gutverdiener beziehungsweise diejenigen, die ihre Interessen vertreten, dann einen Anreiz, ihre Präferenzen für öffentliche Güter zu untertreiben, beziehungsweise Widerstand gegen eine Ausweitung der Bereitstellung öffentlicher Güter zu leisten. Diese Logik kehrt sich allerdings um, wenn optimal gestaltete Einkommensteuern zur Finanzierung herangezogen werden. Dann sind es die Geringverdiener, die die Finanzierungslast tragen, weil eine Ausweitung der Bereitstellung öffentlicher Güter dazu führt, dass die Wohlfahrtskosten umverteilernder Besteuerung steigen. Um eine von diesem strategischen Verhalten unbeeinflusste Kosten-Nutzen-Analyse bei der

21 In Hellwig (2005) wird dieses Thema ausführlicher diskutiert.

Bereitstellung öffentlicher Güter vornehmen zu können, müssen dann sowohl das Steuersystem als auch die Bereitstellungsregel modifiziert werden. Bierbrauer (2014) zeigt für ein Modell nicht-linearer Einkommensbesteuerung, dass dies folgendes Muster erzeugt: Ein Anstieg der Nachfrage nach öffentlichen Gütern führt nicht nur zu höheren Ausgaben für öffentliche Güter sondern auch zu einer Ausweitung der umverteilenden Steuerpolitik. Analog sollte eine niedrige Nachfrage nach öffentlichen Gütern auch zu Einschränkungen bei der Umverteilung führen. Die Geringverdiener verlieren den Anreiz, ihre Nachfrage nach öffentlichen Gütern zu untertreiben, wenn sie fürchten müssen, dass die Transfereinkommen gekürzt werden. Analog werden die Gutverdiener ihre Präferenzen für öffentliche Güter nicht übertreiben, wenn sie wissen, dass eine stärkere Bereitstellung öffentlicher Güter auch mit einer Ausweitung der Verteilungspolitik einhergeht. Diese Analyse liefert also eine mögliche Erklärung dafür, dass in modernen Wohlfahrtsstaaten nicht nur in ausgeprägtem Maß umverteilt wird, sondern dies auch mit hohen Staatsausgaben in anderen Bereichen verbunden ist.

3.5 Optimale Steuern in intertemporaler Perspektive

Ausbildungsentscheidung und Berufswahl. Die Analyse von Mirrlees unterstellt Menschen, deren produktiven Fähigkeiten gegeben sind. Vor diesem Hintergrund werden dann die Effizienzkosten und die Wohlfahrtsgewinne eines umverteilenden Einkommensteuersystems analysiert. Tatsächlich hängen die produktiven Fähigkeiten aber von der gesamten Bildungsbiographie eines Steuerzahlers ab. Wird die Analyse nicht dadurch verfälscht, dass die Wirkung der Besteuerung auf die Anreize, sich auszubilden oder sich auf einen bestimmten Beruf zu spezialisieren, ausgeblendet werden? Natürlich schwächt es die Arbeitsanreize der gutverdienenden Ingenieurinnen, wenn sie Einkommensteuern zahlen müssen. Aber ist nicht das eigentliche Problem, dass niemand mehr Ingenieurin werden will, wenn die finanziellen Vorteile, die mit dem Erwerb dieser Qualifikation verbunden sind, durch das Steuersystem eliminiert werden?

Bovenberg und Jacobs (2005) zeigen, dass diese Vermutung so nicht stimmt. Wenn Menschen zwei für die Einkommenserzielung zentrale Entscheidungen treffen, als junger Mensch die Entscheidung, welche Qualifikation sie erwerben und als älterer, fertig ausgebildeter Mensch, mit wie viel Engagement sie Arbeitseinkommen erzielen, ergeben sich auch zwei Möglichkeiten, mit steuerpolitischen Instrumenten auf diese Entscheidungen einzuwirken. Das Erwerben einer einkommensträchtigen Qualifikation ist dabei nicht nur im eigenen Interesse, sondern geht mit einer sogenannten fiskalischen Externalität einher, da eine Ingenieurin eine ergiebigeren Steuerzahlerin ist als eine technische Zeichnerin. Eine optimal gestaltete Steuerpolitik wird daher mit Subventionen auf die Entscheidung einwirken, ein Universitätsstudium zu beginnen.²² Eine Quantifizierung dieser fiskalischen Externalität anhand von US-

22 Bovenberg und Jacobs unterstellen dabei, dass alle Steuerzahler in gleicher Weise vom Bildungssystem profitieren. Wenn einige mehr profitieren als andere beinhaltet ein optimales Steuersystem dagegen eine Besteuerung des Bildungserwerbs, siehe Hellwig (2008).

amerikanischen Daten findet sich bei Findeisen und Sachs (2015). Die Autoren zeigen, dass, ausgehend vom derzeitigen amerikanischen Steuersystem, die stärkere Subventionierung eines Hochschulstudiums für Kinder einkommensschwacher Familien selbstfinanzierend wäre, d.h. die Kosten einer solchen Subvention würden vollständig von den später anfallenden höheren Einnahmen aus der Einkommensteuer gedeckt.

Eine Reihe von jüngeren Arbeiten hat in den Blick genommen, wie das Steuersystem auf die Entscheidung, einen bestimmten Beruf zu ergreifen, einwirkt, siehe Philippon (2010), Rothschild und Scheuer (2013, 2015) oder Lockwood et al. (2015). Diesen Arbeiten ist gemeinsam, dass sie keine vollständig optimierte Politik betrachten, sondern unterstellen, dass lediglich die Einkommensteuer als Instrument zur Verfügung steht, um sowohl Verteilungsziele zu verfolgen als auch um auf die Berufswahl einzuwirken. Warum aber sollte überhaupt auf die Berufswahl eingewirkt werden? Anders: Warum sollte der Staat überhaupt die Entscheidung einer jungen Frau beeinflussen, ob sie eine Karriere als Ingenieurin oder als Bankerin anstrebt? Philippon (2010) und Rothschild und Scheuer (2015) unterscheiden dazu zwischen Sektoren, in denen nützliche Dinge getan werden, etwa die Entwicklung umweltfreundlicher Technologien, und Sektoren, in denen es im Wesentlichen darauf ankommt, sich auf Kosten anderer zu bereichern. Rothschild und Scheuer verwenden hierfür die Metapher des Goldsuchens. Einem Goldsucher geht es darum, den anderen Goldsuchern zuvorkommen, ohne dass es für den Rest der Welt von Belang wäre, welcher Goldsucher nun der glücklichere ist. Die Finanzindustrie ist potenziell auch noch gefährlich, weil sie immer wieder Krisen produziert, die auch in allen anderen Sektoren Kosten verursachen. Das Steuersystem sollte es demnach für junge Frauen attraktiver machen, Ingenieurin zu werden, und es weniger attraktiv machen, auf eine Karriere in der Finanzindustrie zu zielen.

Dieser Literatur fehlt allerdings eine Strategie zur Quantifizierung der intersektoralen Externalitäten und damit auch die Möglichkeit, den Nachweis, dass mehr Ingenieurinnen besser für das Gesamtsystem wären als mehr Bankerinnen, in überzeugender Weise zu führen. In Ermangelung eines objektiven Maßes intersektoraler Externalitäten behilft sich die Arbeit von Lockwood et al. (2015) mit subjektiven Einschätzungen darüber, wie wichtig unterschiedliche Berufsgruppen für das gesellschaftliche Wohlergehen sind. Wenn man die eigenen Ansichten darüber wie stark Lehrer, Rechtsanwälte, oder Krankenpfleger über- oder unterbezahlt sind in das Modell einspeist, ergibt sich ein aus persönlicher Sicht optimales Einkommensteuersystem. Das Autorenteam hat auch eine App entwickelt, die es jedermann erlaubt dies zu tun.

Die Einkommenssteuer kann auch hilfreich sein um ein vollständig anders gelagertes Externalitätenproblem zu lösen, siehe Corneo (2002), Aronsson und Johansson-Stenman (2008) oder Kanbur und Tuomala (2013). Wenn hoher Konsum mit einem Gewinn an sozialem Status einhergeht, und analog niedriger Konsum mit einem Verlust desselben, dann stellen sich alle Nachbarn schlechter, wenn Herr Jones sich anstrengt, um ein höheres Einkommen zu erzielen. In der Konsequenz würden sich alle besser stellen, wenn sie weniger Einkommen und mehr Freizeit hätten. Wenn alle auf Konsum verzichten, ist dies auch nicht mit einem Verlust an sozialem Status verbunden. Ein progressives Einkommensteuersystems

kann helfen diesen besseren Zustand zu erreichen weil es die Erträge größerer Anstrengungen zur Einkommenserzielung reduziert. Progressive Einkommensbesteuerung ist dann schon aus Effizienzgründen wünschenswert ist, d.h. es sind dann keine spezifischen Verteilungsziele mehr erforderlich, um ein System mit im Einkommen steigenden Grenzsteuersätzen zu rechtfertigen.

New Dynamic Public Finance. Die Literatur zu New Dynamic Public Finance betrachtet eine dynamische Erweiterung des Modells von Mirrlees. Haushalte leben über mehrere Perioden und sind in Bezug auf ihre künftigen produktiven Fähigkeiten mit Unsicherheit konfrontiert.²³ Ausgangspunkt dieser Literatur ist ein Aufsatz von Mikhael Golosov, Narayana Kocherlakota und Aleh Tsyvinski aus dem Jahr 2003. Diese Arbeit markiert sowohl aus methodischer als auch aus inhaltlicher Perspektive einen Wendepunkt in der Literatur zur optimalen Besteuerung von Kapitaleinkommen.

Ehe wir diese Besonderheiten eingehender schildern, sollten wir uns zunächst das zugrundeliegende Allokationsproblem genauer ansehen: Im klassischen Mirrlees-Rahmen leben Haushalte eine Periode. In dieser Periode erzielen sie Einkommen. Ein Teil der hohen Einkommen wird über das Steuersystem an die Bezieher niedriger Einkommen umverteilt. In der dynamischen Erweiterung kommt zu diesem Umverteilungsproblem ein Versicherungsproblem hinzu. Es stellt sich also nicht nur die Frage, wie heute zwischen „arm“ und „reich“ umverteilt werden sollte, sondern beispielsweise auch die Frage, wie ein Haushalt, der heute ein mittleres Einkommen erzielt, vom Steuer- und Transfersystem behandelt werden sollte, wenn er morgen, etwa aufgrund von mehr Berufserfahrung, bessere Möglichkeiten der Einkommenserzielung hat, oder auch wenn er morgen aufgrund eines Autounfalls gar keine Möglichkeiten der Einkommenserzielung mehr hat.

Der methodische Beitrag von Golosov et al. (2003) besteht darin, einen sogenannten „Mechanismus-Design“-Ansatz zu verwenden, um optimale Steuern in einem dynamischen System mit Unsicherheit über künftige Fähigkeiten zu charakterisieren. Ein solcher Ansatz stellt sicher, dass das resultierende Steuersystem auch „wahrhaft“ optimal ist: Die resultierende Allokation kann nicht verbessert werden, solange sowohl die Budgetrestriktion der Ökonomie als Ganzes berücksichtigt werden muss als auch der Umstand, dass Individuen private Information über die Entwicklung ihrer produktiven Fähigkeiten haben. Dies unterscheidet den Mechanism-Design-Ansatz von dem Analyserahmen des Ramsey-Modells, wie ihn etwa Chamley (1986) und Judd (1985) verwenden, um die optimale Besteuerung von Kapitaleinkommen zu analysieren. Wie weiter oben ausgeführt, basiert das Ramsey-Modell auf der Optimierung in einer Menge suboptimaler Politiken. Der von Golosov et al. (2003) verwendete Mechanism-Design-Ansatz ist gegenüber dieser Kritik gefeit.

Das Hauptresultat der Analyse lautet wie folgt: Die Kombination eines Verteilungs- und eines Versicherungsproblems führt dazu, dass die Besteuerung von Kapitaleinkommen wünschens-

23 Werning (2007) behandelt dagegen auch den Fall über mehrere Perioden lebender Haushalte mit über die Zeit nicht veränderlichen produktiven Fähigkeiten.

wert wird.²⁴ Es steht damit im Gegensatz zu den weiter oben besprochenen klassischen Resultaten, die entweder das Ramsey-Modell oder den Analyserahmen von Atkinson und Stiglitz (1976) zugrundelegen und zu dem Ergebnis gelangen, dass Kapitaleinkommen unbesteuert bleiben sollten. Wie kommt das? Betrachten wir wieder den Mittelschichthaushalt, der morgen entweder bessere oder schlechtere Möglichkeiten der Einkommenserzielung hat. Dieser Haushalt möchte Ersparnisse bilden, insbesondere, um für den schlechteren Fall Vorsorge zu treffen. Stellt sich aber der gute Fall ein, sitzt er auf seinen Ersparnissen und hat daher keinen Anreiz, sein gestiegenes Einkommenspotential auszuschöpfen. Ein optimales Steuersystem will diese Situation vermeiden und zielt daher darauf, die Erträge der Ersparnisbildung in genau denjenigen Situationen zu reduzieren, in denen Individuen gute Möglichkeiten der Einkommenserzielung haben. Etwas einfacher formuliert: Die Besteuerung von Kapitaleinkommen ist aus Anreizgründen wünschenswert. Sie soll die produktiven Mitglieder der Gesellschaft davon abhalten, es sich mit ihren Ersparnissen zu bequem zu machen.²⁵

Der Ansatz von Golosov et al. (2003) wurde seither von zahlreichen Autoren weiterentwickelt. Um ein paar Beispiele zu geben: Kocherlakota (2005) stellt die Frage nach den Eigenschaften von konkreten Steuersystemen, welche die Ergebnisse des abstrakteren Mechanismus-Design Ansatzes implementieren. Golosov und Tsyvinski (2006) und Acemoglu et al. (2008) untersuchen, inwiefern auch eine Marktlösung des oben beschriebenen Versicherungsproblems denkbar ist beziehungsweise inwiefern eine Lösung auf der Grundlage des staatlichen Gewaltmonopols erforderlich ist. Farhi und Werning (2012) analysieren die quantitativen Implikationen, also die Frage, wie hoch die Steuersätze auf Kapitaleinkommen ausfallen sollten. Albanesi und Sleet (2006) und Kapicka (2013) verfeinern die mathematischen Methoden zur Analyse des dynamischen Mechanismus-Design-Problems. Wir können den Verästelungen dieser Literatur hier nicht weiter nachgehen. Dass es diese Verästelungen gibt, belegt dass die Arbeit von Golosov et al. (2003) eine zentrale Referenz für die dynamische Analyse optimaler Steuersysteme geworden ist.

Heterogene Präferenzen und die Besteuerung von Kapitalerträgen, Vermögen und Erbschaften. Nachdem wir nun sowohl auf den Erwerb von Qualifikationen durch künftige Steuerzahler geschaut haben als auch auf das Versicherungsbedürfnis aktueller Steuerzahler gegenüber dem Verlust ihrer Arbeitsproduktivität, bleibt uns der Blick auf das im Lauf des Erwerbslebens angehäuften Vermögen und die Besteuerung von Erbschaften.

Wir haben an früherer Stelle gesehen, dass Heterogenität, die über Unterschiede in den produktiven Fähigkeiten hinausgeht, es wünschenswert machen kann, neben der Einkommenssteuer zusätzliche Instrumente der Besteuerung einzusetzen. Zur Erinnerung: Heterogenität in

24 Der zugrundeliegende Formalismus wird häufig als inverse Euler-Gleichung bezeichnet. Die inverse Euler-Gleichung geht zurück auf eine Arbeit von Rogerson (1985). Der Beitrag von Golosov et al. (2003) besteht darin, ihre Implikationen für ein intertemporales Modell optimaler Besteuerung herauszuarbeiten.

25 Entscheidend ist dabei die Unsicherheit über die künftigen produktiven Fähigkeiten. Sind diese nicht unsicher sondern unveränderlich ergibt sich eine intertemporale Erweiterung des Modells von Atkinson und Stiglitz (1976) mit der Konsequenz, dass Kapitaleinkommen nicht besteuert werden sollten, siehe Werning (2007).

den Präferenzen für unterschiedliche Konsumgüter kann zum Einsatz von Verbrauchsteuern führen, siehe die Diskussion zur Arbeit von Atkinson und Stiglitz (1976) in Abschnitt 2.3, und Heterogenität in den Präferenzen für öffentliche Güter kann den Einsatz von Nutzungsgebühren rechtfertigen, siehe Abschnitt 3.4.

Einige Arbeiten der letzten Jahre nutzen einen ähnlichen Ansatz, um die optimale Besteuerung von Kapitaleinkommen zu analysieren. Cremer et al. (2001) betrachten eine optimale Besteuerung von Haushalten, die sich nicht nur im laufenden Einkommen sondern auch im Vermögen unterscheiden, und zeigen, dass eine Vermögensbesteuerung typischerweise wünschenswert ist.²⁶ Diamond und Spinnewijn (2011) treffen die Annahme, dass die Sparneigung einkommensstarker Haushalte größer ist als diejenige einkommensschwacher Haushalte, und zeigen, dass dies eine Besteuerung der Erträge aus Ersparnisbildung wünschenswert macht. Piketty und Saez (2013) analysieren eine optimale Erbschaftsteuer unter der Annahme, dass Individuen sich nicht nur in ihren produktiven Fähigkeiten unterscheiden sondern auch in der Intensität des Wunsches, ihren Nachkommen eine Erbschaft zu hinterlassen.

Wir haben weiter oben im Abschnitt zur Arbeit von Atkinson und Stiglitz (1976) gesehen, dass Modelle optimaler Besteuerung an Plausibilität einbüßen können, wenn der Heterogenität von Präferenzen eine zu starke Bedeutung zukommt. Dieses Thema ist auch hier relevant. Ein gegebenes mathematisches Modell, demzufolge sich Individuen nicht nur in ihren produktiven Fähigkeiten sondern auch noch in einer weiteren Dimension unterscheiden, lässt sich auf unterschiedliche Arten interpretieren. In Modell 1 haben manche Steuerzahler gesunde Eltern, die keiner Pflege bedürfen, und andere Steuerzahler haben pflegebedürftige Eltern. In Modell 2 haben manche Steuerzahler Eltern, die auch im Alter noch gerne verreisen und deshalb nur ein geringes Erbe hinterlassen, während andere großen Wert darauf legen, ihren Nachkommen etwas Gutes zu tun. Der Formalismus beider Modelle kann identisch sein, und doch würde man in der ersten Interpretation eher von Lebensrisiken sprechen, die ein Thema für die Steuerpolitik und die sozialen Sicherungssysteme sind, während die zweite Interpretation das nicht nahelegt. Die Neigung von Rentnern, in Urlaub zu fahren oder Erbschaften zu hinterlassen, ist typischerweise kein Anlass für steuerpolitische Interventionen.

In jedem Fall fällt ins Auge, dass eine Diskussion von Kapitalertragsteuern und Erbschaftsteuern, die auf Heterogenität in den Präferenzen der Sparer oder Erblasser abstellt, wenig mit der Diskussion von Erbschaftsteuern im politischen Prozess zu tun hat. Dort entzündet sich die Kontroverse typischerweise an der Frage, ob eine Erbschaftsteuer die Finanzierungsbedingungen und damit auch die Investitionsmöglichkeiten mittelständischer Unternehmen in Mitleidenschaft ziehen würde.²⁷ Es ist ein Defizit der Literatur zur optimalen Besteuerung von Vermögen und Erbschaften, dass sie zu dieser Diskussion keinen Beitrag leistet. Dies

26 Die Arbeit verwendet den Ansatz von Armstrong und Rochet (1999), um das zweidimensionale Mechanism-Design-Problem zu lösen, das sich ergibt, wenn Haushalte sich sowohl in ihren produktiven Fähigkeiten als auch in ihrem ererbten Vermögen unterscheiden.

27 Ein Überblick über die Argumente in dieser Diskussion findet sich in Brunner (2014).

würde erfordern, die Themen Unternehmensfinanzierung und -kontrolle in die Analyse einzubeziehen. Meines Wissens gibt es hierzu noch keine einschlägige Literatur.²⁸

4 Ausblick: Zurück zu den Grundlagen

Wir haben gesehen, welche Themen der Theorie optimaler Steuern in den letzten Jahren intensiv bearbeitet wurden: Optimale Steuersysteme werden mittlerweile verstärkt kalibriert. Dies erlaubt eine präzisere Beantwortung etwa der Frage, wie hoch der Grenzsteuersatz derer sein sollte, die zum Top-Perzentil in der Einkommensverteilung gehören. Wir haben die verstärkte Auseinandersetzung mit der Frage, ob die Beschäftigung von Geringverdienern subventioniert werden sollte, nachgezeichnet. Eine weitere Erkenntnis ist, dass Steuerwettbewerb nicht nur die Ansiedlung von Firmen betrifft, sondern auch die Frage, wo Menschen leben und arbeiten wollen. Die Literatur der letzten Jahre hat zudem einen Analyserahmen entwickelt, der es nicht nur erlaubt, über optimale Steuerpolitik zu sprechen, sondern auch über die Wechselwirkungen zwischen Steuer- und Ausgabenpolitik. Schließlich gab es eine vertiefte Auseinandersetzung mit der Wirkung der Steuerpolitik über den Lebenszyklus hinweg, angefangen von Ausbildung und Berufswahl, über das Versicherungsbedürfnis von Einkommensbezieher*innen gegen den Verlust ihrer produktiven Fähigkeiten etwa aufgrund eines Unfalls oder aufgrund einer Krankheit, bis hin zum Hinterlassen von Erbschaften.

Wir wollen zum Ende darüber spekulieren, was wohl die Themen eines Überblicks sein könnten, der in 15 Jahren geschrieben wird. Ich vermute, dass dabei die Auseinandersetzungen mit den wohlfahrtsökonomischen Grundlagen eine zentrale Rolle spielen dürfte. Dies umfasst sowohl die Auseinandersetzung mit dem gängigen Verhaltensmodell, das rational handelnde und gut informierte Individuen unterstellt, als auch die Auseinandersetzung mit dem Wohlfahrtsbegriff, also den Werturteilen, die in die Analyse einfließen.

Fangen wir mit dem Verhaltensmodell an. Die Arbeiten von Chetty et al. (2009) und Finkelstein (2009) haben empirisch nachgewiesen, dass die Auffälligkeit oder „Salienz“ der Besteuerung also das Ausmaß, in dem Individuen bei Ihren Entscheidungen die steuerlichen Implikationen vor Augen haben, einen Einfluss auf das Verhalten der Steuerzahler hat. Beispielsweise gibt es stärkere Reaktionen auf Verbrauchsteuern, wenn diese deutlich auf Preisschildern ausgewiesen werden. Einige Autoren haben dies zum Anlass genommen, die Salienz der Besteuerung in die Analyse optimaler Steuersysteme einzubeziehen, siehe etwa Goldin (2015) oder Farhi und Gabaix (2015). Diese Arbeiten sind Beispiele für das Aufgreifen von verhaltensökonomischer Forschung in der Theorie optimaler Besteuerung. Vermutlich werden wir in den nächsten Jahren weitere Beispiele sehen.

28 Die Analyse von Grossman und Strulik (2010) untersucht welche Effizienzwirkungen eine Verschonung von Betriebsvermögen für Familienbetriebe hat. Die Analyse enthält jedoch keine mikrofundierte Modellierung von Kapitalmarktfriktionen.

Die meisten Arbeiten zur Theorie optimaler Besteuerung basieren auf einem utilitaristischen Wohlfahrtsmaß. Dieses Vorgehen ist von verschiedenen Autoren kritisiert worden. Beispielsweise zeigen Mankiw und Weinzierl (2010), dass dieser Ansatz, wenn er konsequent zu Ende gedacht würde, eine Besteuerung nicht nur nach dem erzielten Einkommen sondern auch nach der Körpergröße erfordern würde, weil große Menschen häufiger über hohe produktive Fähigkeiten verfügen. Saez und Stantcheva (2016) argumentieren, dass der utilitaristische Ansatz unverhältnismäßig viel Umverteilung generiert. Die Arbeiten von Weinzierl (2014a,b) und Saez und Stantcheva (2016) schlagen daher alternative Wohlfahrtsmaße vor, liefern jedoch noch keine vollständige Charakterisierung optimaler Einkommensteuern auf dieser Grundlage. Daher lässt sich vermuten, dass weitere Arbeiten folgen werden.

Selbst wenn sich die Implikationen alternativer Wohlfahrtsmaße für eine optimale Steuerpolitik durch künftige Forschung klären lassen, bleibt aber die grundsätzliche Frage nach der politischen Relevanz wohlfahrtsmaximierender Ansätze offen. Die Arbeiten von Bierbrauer und Boyer (2016), Brett und Weymark (2012) und Zoutman et al. (2015) zeigen sowohl theoretisch als auch empirisch, dass die Vorstellung, wonach sich eine Gesellschaft im politischen Prozess auf das relevante Wohlfahrtsmaß verständigt, unangemessen ist. Im politischen Prozess werden Partikularinteressen und die Interessen der Mittelschicht in einem Maß berücksichtigt, die jeder Vorstellung von Wohlfahrtsmaximierung zuwider laufen. Dies legt ein Forschungsprogramm nahe, das bei der Betrachtung der Steuerpolitik nicht nur die Verhaltensreaktionen der Steuerzahler sondern auch die Auswirkungen in der politischen Auseinandersetzung einbezieht. Wenn ich einen Wunsch frei hätte, dann den nach einer Theorie, die es erlaubt, die Menge aller Steuerreformen zu charakterisieren, die sowohl wohlfahrtssteigernd als auch politisch realisierbar sind.

Darüber hinaus bleibt natürlich die Hoffnung, dass die Charakterisierung wohlfahrtsmaximierender Steuerpolitik, etwa auf dem Weg über die akademische Politikberatung, vielleicht doch eine positive Wirkung entfaltet.

Literatur

Acemoglu, D., M. Golosov, und A. Tsyvinski (2008), Markets versus governments, *Journal of Monetary Economics* 55, S. 159-189.

Akerlof, G.A. (1978), The economics of “tagging” as applied to the optimal income tax, welfare programs, and manpower planning, *American Economic Review* 68(1), S. 8-19.

Albanesi, S. und C. Sleet (2006), Dynamic optimal taxation with private information, *Review of Economic Studies* 73(1):1-30.

Armstrong, A. und J. Rochet (1999). Multidimensional screening: A user's guide, *European Economic Review* 43, S. 959-979.

- Aronsson, T. und Johannson-Stenman, O. (2008). When the Joneses' consumption hurts: Optimal public good provision and nonlinear income taxation. *Journal of Public Economics*, 92. S. 986-997.
- Atkeson, A., V. Chari und P. Kehoe (1999), Taxing capital income: A bad idea, *Federal Reserve Bank of Minneapolis Quarterly Review* 23, S. 3-17.
- Atkinson, A. und J. Stiglitz (1976), The design of tax structure: Direct versus indirect taxation, *Journal of Public Economics* 1, S. 55-75.
- Bach, S., G. Corneo, und V. Steiner (2012). Optimal top marginal tax rates under income splitting for couples, *European Economic Review* 56(6), S. 1055-1069.
- Bierbrauer, F. (2009), Optimal income taxation and public good provision with endogenous interest groups, *Journal of Public Economic Theory* 11, S. 311-342.
- Bierbrauer, F. (2011a), On the optimality of optimal income taxation, *Journal of Economic Theory* 146, S. 2105-2116.
- Bierbrauer, F. und P. Boyer (2014), The Pareto-frontier in a simple Mirrlees model of income taxation, *Annals of Economics and Statistics* 113/14, S. 185-206.
- Bierbrauer, F., C. Brett und J. Weymark (2013), Strategic nonlinear income tax competition with perfect labor mobility, *Games and Economic Behavior* 82(0), S. 292-311.
- Bierbrauer, F. (2011b), Incomplete contracts and excludable public goods, *Journal of Public Economics* 95 (7-8), S. 553-569.
- Bierbrauer, F. (2012), Distortionary taxation and the free-rider problem, *International Tax and Public Finance* 19, S. 732-752.
- Bierbrauer, F. (2014), Optimal tax and expenditure policy with aggregate uncertainty, *American Economic Journal: Microeconomics* 6, S. 205-257.
- Bierbrauer, F. und P. Boyer (2016), Efficiency, welfare and political competition, Forthcoming, *Quarterly Journal of Economics*.
- Brunner, J. (2014), Die Erbschaftssteuer – Bestandteil eines optimalen Steuersystems?, *Perspektiven der Wirtschaftspolitik* 10, S. 199-218.
- Blundell, R., A. Bozio und G. Laroque, (2011), Labor supply and the extensive margin, *American Economic Review* 101(3), S. 482-486.
- Boadway, R. und M. Keen, (1993), Public goods, self-selection and optimal income taxation, *International Economic Review* 34, S. 463-478.

- Boiteux, M. (1956), Sur la gestion des Monopoles Publics astreints à l'équilibre budgétaire, *Econometrica* 24, S. 22-40.
- Bovenberg, L. und B. Jacobs (2005), Redistribution and education subsidies are Siamese twins, *Journal of Public Economics* 89(11-12), S. 2005-2035.
- Brennan, G. und J. Buchanan (1980), The Power to Tax: Analytical Foundations of a Fiscal Constitution, Indianapolis, In: Liberty Fund, Inc.
- Brett, C. und J. Weymark (2012), Citizen candidates and voting over incentive compatible nonlinear income tax schedules, Unveröffentlichtes Diskussionspapier.
- Chamley, C. (1986), Optimal taxation of capital income in general equilibrium with infinite lives, *Econometrica* 54, S. 607-622.
- Chetty, R., A. Looney und K. Kroft (2009), Salience and taxation: Theory and evidence, *American Economic Review* 99(4), S. 1145-77.
- Choné, P. und G. Laroque (2010), Negative marginal tax rates and heterogeneity, *American Economic Review* 100(5), S. 2532-47.
- Choné, P. und G. Laroque (2011), Optimal taxation in the extensive model, *Journal of Economic Theory* 146(2) S. 425-453.
- Corneo, G. (2002). The efficient side of progressive income taxation. *European Economic Review* 46, S. 1359-1368.
- Cremer, H., P. Pestieau und J. Rochet (2001), Direct versus indirect taxation: The design of the tax structure revisited, *International Economic Review* 42 S. 781-799.
- Diamond, P. (1975), A many-person Ramsey tax rule, *Journal of Public Economics* 4, S. 335-342.
- Diamond, P. (1998), Optimal income taxation: An example with a u-shaped pattern of optimal marginal tax rates, *American Economic Review* 88, S. 83-95.
- Diamond, P. und E. Saez, (2011), The case for a progressive tax: From basic research to policy recommendations, *Journal of Economic Perspectives* 25(4), S. 165-90.
- Diamond, P. und J. Spinnewijn (2011), Capital income taxes with heterogeneous discount rates, *American Economic Journal: Economic Policy* 3(4), S. 52-76.
- Farhi, E. und X. Gabaix (2015), Optimal taxation with behavioral agents, unveröffentlichtes Diskussionspapier.
- Farhi, E. und I. Werning (2012), Capital taxation: Quantitative explorations of the inverse Euler equation, *Journal of Political Economy* 120, S. 398-445.

- Findeisen, S. und D. Sachs (2015), Designing efficient college and tax policies, unveröffentlichtes Diskussionspapier.
- Finkelstein, A. (2009), EZ-tax: Tax salience and tax rates, *Quarterly Journal of Economics* 124(3), S. 969-1010.
- Fleurbaey, M. und F. Maniquet (2006), Fair income tax, *Review of Economic Studies* 73(1), S. 55-83.
- Gaube, T. (2000), When do distortionary taxes reduce the optimal supply of public goods, *Journal of Public Economics* 76 S. 151-180.
- Goldin, J. (2015), Optimal tax salience. Forthcoming, *Journal of Public Economics*.
- Golosov, M., N. Kocherlakota, und A. Tsyvinski (2003), Optimal indirect and capital taxation, *Review of Economic Studies* 70, S. 569-587.
- Golosov, M. und A. Tsyvinski (2006), Designing optimal disability insurance: A case for asset testing, *Journal of Political Economy* 114, S. 257-269.
- Guesnerie, R. (1995), A Contribution to the Pure Theory of Taxation. *Cambridge University Press*.
- Grossman, V. und H. Strulik (2010), Should continued family firms face lower taxes than other estates?, *Journal of Public Economics* 94, S. 87-101.
- Hammond, P. (1979), Straightforward individual incentive compatibility in large economies, *Review of Economic Studies* 46, S. 263-282.
- Hansen, E. (2015), On the ambiguous sign of the optimal utilitarian marginal income tax, unveröffentlichtes Diskussionspapier.
- Hellwig, M. (2004), Optimal income taxation, public goods provision and public-sector pricing: A contribution to the foundations of public economics, Preprint 2004/14, Max Planck Institute for Research on Collective Goods, Bonn.
- Hellwig, M. (2005), A utilitarian approach to the provision and pricing of excludable public goods, *Journal of Public Economics* 89, S. 1981-2003.
- Hellwig, M. (2007a), A contribution to the theory of optimal utilitarian income taxation, *Journal of Public Economics* 91, S. 1449-1477.
- Hellwig, M. (2007b), The provision and pricing of excludable public goods: Ramsey-Boiteux pricing versus bundling, *Journal of Public Economics* 91, S. 511-540.

- Hellwig, M. (2008), A Generalization of the Atkinson-Stiglitz (1976) Theorem on the Undesirability of Nonuniform Excise Taxation, Preprint 2008/47, Max Planck Institute for Research on Collective Goods, Bonn.
- Jacquet, L., E. Lehmann und Van der Linden, B. (2013), Optimal redistributive taxation with both extensive and intensive responses, *Journal of Economic Theory* 148(5), S. 1770-1805.
- Jehiel, P. und L. Lamy, (2015), A mechanism design approach to the Tiebout hypothesis, unveröffentlichtes Diskussionspapier.
- Judd, K. (1985), Redistributive taxation in a simple perfect foresight model, *Journal of Public Economics* 28, S. 59-83.
- Kanbur, R. und Tuomala, M. (2013). Relativity, Inequality and Optimal Nonlinear Income Taxation. *International Economic Review*, 54. S. 1199-1217.
- Kapicka, M. (2013), Efficient allocations in dynamic private information economies with persistent shocks: A first-order approach, *Review of Economic Studies* 80(3), S. 1027-1054.
- Kaplow, L. (2006), On the undesirability of commodity taxation even when income taxation is not optimal, *Journal of Public Economics* 90, S. 1235-1250.
- Kleven, H., C. Kreiner und E. Saez (2009), The optimal income taxation of couples. *Econometrica* 77(2), S. 537-560.
- Kleven, H., C. Landais und E. Saez (2013), Taxation and international migration of superstars: Evidence from the European football market, *American Economic Review* 103(5), S. 1892-1924.
- Kleven, H., C. Landais, E. Saez und E. Schultz (2014), Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in Denmark, *Quarterly Journal of Economics* 129(1), S. 333-378.
- Kocherlakota, N. (2005), Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation, *Econometrica* 73, S. 1587-1621.
- Laroque, G. (2005), Indirect taxation is harmful under separability and taste homogeneity: A simple proof, *Economic Letters* 87, S. 141-144.
- Lehmann, E., L. Simula und A. Trannoy (2014), Tax me if you can! Optimal nonlinear income tax between competing governments, *Quarterly Journal of Economics* 129(4), S. 1995-2030.
- Lindahl, E. (1919), *Die Gerechtigkeit der Besteuerung*. Lund.

- Lockwood, B., C. Nathanson und G. Weyl (2015), Taxation and the allocation of talent, unveröffentlichtes Diskussionspapier.
- Lockwood, B. und M. Weinzierl (2015), De gustibus non est taxandum: Heterogeneity in preferences and optimal redistribution, *Journal of Public Economics* 124, S. 74-80.
- Mankiw, G. und M. Weinzierl (2010), The optimal taxation of height: A case study of utilitarian income redistribution, *American Economic Journal: Economic Policy* 2(1), S. 155-76.
- Mankiw, N. G., M. Weinzierl und D. Yagan (2009), Optimal taxation in theory and practice, *Journal of Economic Perspectives* 23(4), S. 147-74.
- Mirrlees, J. (1971), An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, S. 175-208.
- Nava, M., F. Schroyen und M. Marchand (1996), Optimal fiscal and public expenditure policy in a two-class economy, *Journal of Public Economics* 61, S. 119-137.
- Philippon, T. (2010), Financiers vs. engineers: Should the financial sector be taxed or subsidized? *American Economic Journal: Macroeconomics* 2, S. 158-82.
- Piketty, T. (2014), *Capital in the Twenty-First Century*, Belknap Press of Harvard University Press.
- Piketty, T. und E. Saez (2013), A theory of optimal inheritance taxation, *Econometrica* 81(5), S. 1851-1886.
- Ramsey, F. (1927), A contribution to the theory of taxation, *Economic Journal* 37, S. 47-61.
- Rogerson, W. (1985), Repeated moral hazard. *Econometrica* 53, S. 69-76.
- Rothschild, C. und F. Scheuer, (2013), Redistributive taxation in the Roy model, *The Quarterly Journal of Economics* 128(2), S. 623-668.
- Rothschild, C. und F. Scheuer (2015), Optimal taxation with rent-seeking, conditionally accepted, *Review of Economic Studies*.
- Saez, E. (2001), Using elasticities to derive optimal income tax rates. *Review of Economic Studies* 68, S. 205-229.
- Saez, E. (2002), Optimal income transfer programs: Intensive versus extensive labor supply responses, *Quarterly Journal of Economics* 117(3), S. 1039-1073.
- Saez, E. und S. Stantcheva (2016), Generalized social marginal welfare weights for optimal tax theory, *American Economic Review* 106(1), S. 24-45.

- Samuelson, P. (1954), The pure theory of public expenditure, *Review of Economics and Statistics* 36, S. 387-389.
- Scheuer, F. und I. Werning (2015), The taxation of superstars. Unveröffentlichtes Diskussionspapier.
- Schulte, B. (2014), Autonomie für Länder und Kommunen bei der Einkommenssteuer, *Argumente zu Marktwirtschaft und Politik* 126, Stiftung Marktwirtschaft.
- Stiglitz, J. (1982), Self-selection and Pareto-efficient taxation, *Journal of Public Economics* 17, S. 213-240.
- Straub, L. und I. Werning (2015), Positive long run capital taxation: Chamley-Judd revisited, unveröffentlichtes Diskussionspapier.
- Tiebout, C. M. (1956), A pure theory of local expenditures, *Journal of Political Economy* 64(5), S. 416-424.
- Weinzierl, M. (2014a), The promise of positive optimal taxation: normative diversity and a role for equal sacrifice, *Journal of Public Economics* 118, S. 128-142.
- Weinzierl, M. (2014b), Revisiting the classical view of benefit-based taxation, *Harvard Business School*, Unveröffentlichtes Diskussionspapier No. 14-101.
- Werning, I. (2007), Optimal fiscal policy with redistribution, *Quarterly Journal of Economics* 122(3), S. 925-967.
- Zoutman, F., B. Jacobs und E. Jongen (2015), Redistributive politics and the tyranny of the middle class, unveröffentlichtes Diskussionspapier, Erasmus University Rotterdam.