INSTITUT FÜR
VOLKSWIRTSCHAFTSLEHRE
JOHANNES KEPLER UNIVERSITÄT LINZ

Optimale direkte und indirekte Steuern bei unterschiedlicher Anfangsausstattung

von

Johann K. Brunner

Arbeitspapier Nr. 0310
September 2003

Johannes Kepler Universität Linz
Institut für Volkswirtschaftslehre
Altenberger Straße 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

johann.brunner@jku.at
Tel. +43 (0)70 2468 -8248, -9821 (Fax)

„There is no simple route between the Mirrlees model and policy implications for annual income taxes ....”

(Diamond 1998, p.93)

A. Einleitung

Eines der zentralen wirtschaftspolitischen Anliegen ist die „richtige“ Ausgestaltung des Steuersystems. Daher haben Ökonominnen und Ökonomen, die ja nicht nur das Bestehende erklären sondern auch Vorschläge für Verbesserungen machen wollen, sich schon immer dieses Themas angenommen und normative Aussagen dazu getätigt - welche mehr oder weniger in theoretischen oder empirischen Erkenntnissen begründet waren. In den letzten Jahrzehnten hat sich, vor allem unter dem Einfluss der angelsächsischen Ökonomie, auch in der Finanzwissenschaft die modelltheoretische Methode, also die Begründung theoretischer Aussagen durch ihre Herleitung aus mathematischen Modellen, weitgehend durchgesetzt. Ein wichtiges und typisches Beispiel dafür ist eben der Versuch, „optimale Steuern“ durch die Maximierung einer Zielfunktion unter Nebenbedingungen zu finden.

Allerdings ist dieses Unternehmen, also die Theorie der optimalen Besteuerung, durchaus umstritten. Es gibt vermutlich eine große Anzahl von Ökonominnen und Ökonomen, die sie als vielleicht rein mathematisch interessante, aber ziemlich weltfremde Angelegenheit betrachten.[1] Andererseits findet man auch in der Gegenwart immer wieder Beiträge von sehr bekannten Wissenschaftern zu dieser Theorie, was dokumentiert, dass wenigstens für manche das Interesse daran ungebrochen ist.[2] Als ein genereller Einwand gegen die Kritiker kann vorgebracht werden, dass es nicht von vornherein einsichtig ist, warum der Vorwurf der Weltfremdheit gerade bei dieser Theorie berechtigter sein soll als in anderen Bereichen der Volkswirtschaftslehre. Aber selbst wenn man ihn bis zu einem gewissen Grad akzeptiert, kann es gute Gründe geben, sich mit dieser Theorie zu beschäftigen:

Erstens liegt es eben tatsächlich nahe, die Suche nach einem richtigen (optimalen) Steuersystem, die ja – wie erwähnt - immer schon ein Anliegen der Finanzwissenschaft war, auf theoretischer Ebene anhand der ökonomischen Standardmodelle voranzutreiben. Diese Modelle gehören heute zum Werkzeugkasten der ökonomischen Forschung und werden für verschiedene Fragen angewendet, warum dann nicht auch für die Frage nach der Besteuerung? So erscheint es

[1] Das drückt etwa die Antwort eines Kollegen auf die Bemerkung des Autors, dass er in Lehre auch die Theorie der optimalen Besteuerung (kurz) behandelt, aus: „Macht das immer noch jemand?“

[2] So sind in der jüngeren Zeit wieder zwei Überblicksartikel dazu erschienen (Chari und Kehoe 1999, Auerbach und Hines 2002), die zeigen, dass jedenfalls bei der Herausgabe von Handbüchern die Behandlung der Theorie optimaler Steuern nicht fehlen darf.

konsequent, sich von der Modellierung der Beziehung zwischen Effizienz und Gerechtigkeit (Umverteilung), die ja im Zentrum der Steuertheorie steht, einen wichtigen Schritt zu einem besseren Verständnis der grundlegenden Zusammenhänge zu erhoffen.

Zweitens hat die Theorie optimaler Steuern eine Verbreitung erlangt, die ein bloßes Ignorieren als eine problematische Haltung erscheinen lässt, sondern eher eine Auseinandersetzung damit erfordert. Dies umso mehr, als bestimmte Ergebnisse von manchen ihrer Vertreter durchaus als adäquate Unterlegung entsprechender politischer Forderungen angesehen werden.[3] Außerdem zeigt sich, wenn man die historische Entwicklung der Steuern betrachtet, seit etwa den achtziger Jahren des vorigen Jahrhunderts eine deutliche Tendenz zur Senkung der Grenzsteuersätze. Der höchste Marginalsteuersatz lag etwa in den USA Ende der siebziger Jahre noch bei 70 Prozent – in der unmittelbaren Nachkriegszeit sogar bei 91 Prozent - und wurde dann auf etwa 40 Prozent gesenkt. Ähnlich in Großbritannien, wo es eine Senkung von einem Ausgangswert von über 80 Prozent auf etwa 40 Prozent gab, und in Schweden (von mehr als 80 auf unter 40 Prozent).[4] In Deutschland und Österreich gab es nie so extrem hohe Grenzsteuersätze, aber auch hier wurden sie gesenkt, in Deutschland von 56 zunächst auf 53 Prozent und nun auf unter 50, in Österreich im Jahr 1988 von 62 auf 50 Prozent (für Unselbständige beträgt er effektiv nur etwa 44 Prozent). In dieser Entwicklung spiegelt sich offensichtlich eine geänderte Sichtweise wider, die den Schwerpunkt eher auf die Problematik negativer Anreizeffekte legt als auf die Möglichkeiten der Umverteilung. Die Ursachen für diese Verlagerung sind sicherlich vielfältig, aber man kann jedenfalls konstatieren, dass sie parallel mit der Rezeption der entsprechenden Resultate aus der Optimalsteuertheorie, v.a. im Rahmen des Modells von Mirrlees (1971), geht. Für eine Einschätzung dieser Entwicklung erscheint die Erörterung der Struktur der dahinter stehenden Modelle und der Annahmen, auf denen sie aufbauen, nützlich.

Drittens erlaubt es die Kenntnis der Modelle, Änderungen oder Erweiterungen der dahinter stehenden Annahmen zu erörtern, die eine Abbildung anderer - ebenso wichtiger, aber in den Standardmodellen vernachlässigter - Zusammenhänge der realen Welt ermöglichen. Ökonomisches Denken ist heute weitgehend ein „Denken in Modellen", was die Gefahr mit sich bringt, dass alle Aspekte, die nicht modellhaft formuliert werden (können), unberücksichtigt bleiben.

Viertens – als zugestanden außerökonomischer Grund – mag es sein, dass die Struktur des im einleitenden Zitat angesprochenen Mirrlees-Modells tatsächlich einen gewissen Reiz auf formal

[3] Als hervorstechendes Beispiel dafür sei der Titel der Arbeit von Atkeson, Chari und Kehoe (1999) angeführt: „Taxing capital income: a bad idea".

[4] Diese Angaben sollen nur die Größenordnungen andeuten. In manchen Ländern kommt zur Bundeseinkommensteuer noch eine lokale hinzu, die nicht einheitlich ist. Daher ist eine allgemeine Angabe schwierig.

interessierte Ökonomen ausübt; ein Phänomen, das sicher auch in anderen Bereichen der Volkswirtschaftslehre auftritt.

In gewisser Weise kann man die Bemühungen mancher Ökonomen etwa um 1900, einen optimalen Steuertarif aus den Opferprinzipien herzuleiten, als Vorläufer der Theorie der optimalen Besteuerung ansehen (u.a. Cohen-Stuart 1889, Edgeworth 1897, Frisch 1932)[5]. Dabei wurde in mathematisch durchaus komplexen Erörterungen in erster Linie die Frage untersucht, ob mit dem Prinzip des gleichen proportionalen Opfers ein progressiver Einkommensteuertarif begründet werden kann. Dies lief offensichtlich auf die Erörterung des Verlaufs (der Krümmung) der Nutzenfunktion in Abhängigkeit vom Einkommen hinaus – eine Fragestellung, auf die sich vermutlich kaum eine befriedigende Antwort finden lässt. Trotzdem ist die Untersuchung, mit welchen Eigenschaften dieser Nutzenfunktion die Progressivität, Proportionalität oder Regressivität des Steuertarifs gemäß dem gleichen proportionalen Opfer zusammenhängt, interessant.

Der entscheidende Unterschied, der die moderne Theorie der optimalen Besteuerung charakterisiert, ist die Einbeziehung der Verzerrung der Arbeitsangebotsentscheidung. Dies gilt auch für das grundlegende Modell von Ramsey (1927), das zwar üblicherweise mit der Suche nach der optimalen indirekten Besteuerung assoziiert wird, aber diese wird eben erst zu einem Problem, wenn der Haushalt über die Höhe des Arbeitsangebots und damit des Bruttoeinkommens entscheidet; ansonsten ist klarerweise ein einheitlicher Steuersatz auf alle Güter optimal, weil er wie eine Pauschalsteuer wirkt.

Das Mirrlees-Modell kann unmittelbar als eine Erweiterung der Suche nach dem optimalen Steuertarif um die Berücksichtigung des Anreizeffektes beim Arbeitsangebot gesehen werden. Allerdings geht es von einer geänderten Zielvorstellung aus: der gesuchte Tarif soll nicht mehr dem Prinzip einer proportionalen Verringerung der Nutzenpositionen (möglichst) gerecht werden, sondern eine utilitaristische soziale Wohlfahrtsfunktion maximieren. Negative Steuern sind im Mirrlees-Modell nicht ausgeschlossen, so dass optimale Umverteilung Teil der Fragestellung wird – die Lösung stellt ein kombiniertes Steuer-Transfer-Schema dar.[6] Man kann hinter dem Wechsel der Zielvorstellung eine grundsätzlich andere Orientierung sehen: Während hinter den Opferprinzipien die Idee steht, dass das selbst erwirtschaftete Bruttoeinkommen einen Referenzpunkt darstellt, der festlegt, worauf eine Person im Prinzip Anspruch hat, auch wenn ihr in Form von Steuern ein (adäquater) Teil entzogen wird, steht im jüngeren Modell das aggregierte

[5] Zu einer Diskussion siehe u. a. Brunner 1989.
[6] Vertreter des Prinzips der gleichen proportionalen Opfers hatten dagegen die Idee, dass zunächst (mit nicht näher formal beschriebenem Ziel) eine gewünschte Verteilung der Bruttoeinkommen hergestellt werden sollte. Letztere sollten dann der Besteuerung entsprechend dem Opferprinzip unterzogen werden. Siehe etwa Haller 1981.

Einkommen zur Maximierung der sozialen Wohlfahrtsfunktion zur Verfügung.[7] Das selbst erwirtschaftete Bruttoeinkommen hat keine normative Rolle.

Neben der Wirkung der Besteuerung auf die Arbeitsangebotsentscheidung stand schon seit längerer Zeit auch deren Effekt auf die Kapitalbildung im Mittelpunkt vieler Diskussionen. Demgemäß wurde in der Theorie der optimalen Besteuerung auch eine Reihe von Aussagen zur Rolle einer Steuer auf Kapitaleinkommen hergeleitet. Soweit sie aus intertemporalen Erweiterungen der Ansätze von Ramsey und Mirrlees zu Modellen überlappender Generationen entstanden sind, beruhen diese Aussagen de facto auf dem Zusammenhang zwischen der Kapitalbildung und der Arbeitsangebotsentscheidung, wie gerade angemerkt. Zu diesem Zusammenhang können – das ist die Schwierigkeit - kaum verlässliche empirische Aussagen gefunden werden. Dagegen tritt in Modellen mit unendlichem Zeithorizont der Individuen ein anderer Aspekt, nämlich die gegen unendlich gehende spezifische Verzerrung der Sparentscheidung, in den Vordergrund und gibt den Ausschlag für das Resultat, das eine Kapitaleinkommensbesteuerung als wohlfahrtsmindernd charakterisiert.

Die Intention des vorliegenden Beitrages ist, eine kritische Betrachtung der in der Theorie der optimalen Besteuerung angewendeten Grundmodelle und einiger Erweiterungen, die in der jüngeren Vergangenheit entwickelt wurden, vorzunehmen. Naturgemäß kann dieses Vorhaben nur unter Beschränkungen verfolgt werden und so bleiben dabei eine Reihe von Bereichen, die ebenfalls wichtige Ergebnisse für die optimale Gestaltung von Steuern geliefert haben, außer Acht, etwa alle Überlegungen zur Besteuerung von Unternehmen oder die Resultate von Simulationsmodellen zu den Wirkungen von Steuern, ebenso werden die Modelle mit endogenem Wachstum ausgeklammert. Der folgende Abschnitt B enthält eine Darstellung der beiden Grundmodelle von Ramsey und Mirrlees sowie einige Ergänzungen und Anmerkungen dazu. Im Abschnitt C wird auf dynamische Modelle und damit vor allem auf die Frage der Besteuerung von Kapitaleinkommen eingegangen, der letzte Abschnitt D enthält einige Schlussbetrachtungen. Generell wurde versucht, die Grundstruktur der Modelle in möglichst einfacher Form zu präsentieren, um die wesentlichen Annahmen und Eigenschaften deutlich herauszuarbeiten; auf komplexe mathematische Herleitungen wurde dagegen verzichtet.

[7] Die Existenz einer Einkommensteuerfunktion, die dem Prinzips der gleichen proportionalen Opfers unter Einbeziehung der Arbeitsangebotsentscheidung genügt, wurde von Berliant und Gouveia (1993) untersucht.

B. Statische Grundmodelle der optimalen Besteuerung

B.I. Das Ramsey Modell

Wir beginnen mit einer kurzen Darstellung der ersten und für viele weitere Arbeiten grundlegenden Analyse, wie ein Steuersystem konstruiert sein sollte, das ein gegebenes Steueraufkommen für den Staat auf eine solche Weise erbringt, dass ein möglichst geringer Effizienzverlust entsteht. Es sei ein repräsentativer Haushalt gegeben, der bei einem Lohnsatz w ein Arbeitseinkommen wL erzielt (dabei ist L die Arbeitszeit), das er für den Konsum zweier Güter (deren Mengen mit c_1 und c_2 bezeichnet werden) verwenden kann. p_1, p_2 seien die Produzentenpreise der beiden Güter, und zur Vereinfachung nehmen wir an, dass diese Preise sowie w durch Steuern unverändert bleiben (das bedeutet, es wird mit einer linearen Technologie - konstante Grenzproduktivität der Arbeit - produziert).

Der Staat kann proportionale Steuern τ_1, τ_2 auf die beiden Güter sowie τ_w auf Arbeitseinkommen einheben. Die Budgetbedingung des Haushalts lautet somit

$$p_1(1 + \tau_1)c_1 + p_2(1 + \tau_2)c_2 \leq w(1 - \tau_w)L. \qquad \text{(B.I.1)}$$

Man sieht unmittelbar, dass in diesem Modell (wie man es aus jenem mit fixem Budget schließen könnte) ein gleicher Steuersatz auf beide Güter und auf Arbeitseinkommen nicht wie eine Pauschalbesteuerung wirkt, wegen des unterschiedlichen Vorzeichens.[8] Weiters folgt aus (B.I.1) unmittelbar, dass sich die Budgetmenge nicht ändert, wenn eine Steuer weggelassen wird und dafür die beiden anderen angepasst werden. Zum Beispiel kann man die Ungleichung durch $1 - \tau_w$ dividieren und $(1+\tau_i)/(1-\tau_w)$ in der Form $1+\tau_i'$ schreiben, wobei $\tau_i' \equiv (\tau_i + \tau_w)/(1-\tau_w)$ gewählt wird. Die Budgetbedingung (B.I.1) mit drei Steuersätzen τ_1, τ_2, τ_w ist dann äquivalent zur Bedingung $p_1(1+\tau_1') + p_2(1+\tau_2') \leq wL$, in der die Zahl der Steuersätze auf zwei reduziert ist. Daher wird im Folgenden nur mehr ein Steuersystem bestehend aus den beiden Gütersteuern betrachtet; außerdem formulieren wir das Modell in der Weise um, dass optimale Mengensteuern $\hat{\tau}_1, \hat{\tau}_2$ statt der Steuersätze τ_1, τ_2 gesucht werden.[9] Die Konsumentenpreise sind dann $q_i = p_i + \hat{\tau}_i$.

---

[8] Eine Pauschalsteuer würde sich ergeben, wenn der Wert der Freizeit statt des Arbeitseinkommens besteuert werden könnte. Dann käme auf der linken Seite von (B.I.1) der Term $\tau_w wH$ hinzu (H bezeichnet die Freizeit), die rechte Seite würde zu wL, was insgesamt zur Budgetbedingung $p_1(1 + \tau_1)c_1 + p_2(1 + \tau_2)c_2 + w(1 + \tau_w)H \leq w\bar{L}$ führt (mit $\bar{L}$ als exogen gegebener verfügbarer Zeit und $L = \bar{L} - H$) und man sieht, dass sich im Fall $\tau_1 = \tau_2 = \tau_w$ eine Pauschalsteuer ergibt.

[9] Offensichtlich kann man mithilfe der Beziehung $p_i + \hat{\tau}_i = p_i(1+\tau_i)$ stets $\hat{\tau}_i$ aus τ_i errechnen und umgekehrt.

Sei $U(c_1,c_2,L)$ eine quasikonkave Nutzenfunktion, die die Präferenzordnung des Haushalts beschreibt (wobei $\partial U/\partial c_i > 0$ und $\partial U/\partial L < 0$ gilt) und $V(p_1+\hat{\tau}_1, p_2+\hat{\tau}_2, w, x)$ die zugehörige indirekte Nutzenfunktion, wobei x ein eventuell vorhandenes exogenes Einkommen (im Standardfall als null angenommen) bezeichnet. Das Problem des Staates ist, ein gegebenes Steueraufkommen g mithilfe der beiden Gütersteuern zu erzielen, wobei der Effizienzverlust, den der repräsentative Haushalt erleidet, möglichst gering sein soll. Das Optimierungsproblem lautet daher:

$$\max_{\hat{\tau}_1,\hat{\tau}_2} V(p_1+\hat{\tau}_1, p_2+\hat{\tau}_2, w, x), \tag{B.I.2}$$

$$\text{s.t.} \quad \sum_{i=1}^{2} \hat{\tau}_i c_i(p_1+\hat{\tau}_1, p_2+\hat{\tau}_2, w, x) \geq g. \tag{B.I.3}$$

Dabei sind die $c_i(\cdot)$ die Nachfragefunktionen nach den beiden Gütern.

Aus den Bedingungen erster Ordnung für (B.I.2), (B.I.3) ergibt sich nach einigen Umformungen[10] das bekannte Gleichungssystem:

$$\sum_{i=1}^{2} \tau_i \partial c_k^c / \partial \hat{\tau}_i = -\theta c_k, \qquad k = 1, 2. \tag{B.I.4}$$

Dabei ist θ eine Konstante und $c_k^c(p_1+\hat{\tau}_1, p_2+\hat{\tau}_2, w, U)$ die kompensierte Nachfragefunktion nach Gut k, wobei für U das Nutzenniveau im Optimum einzusetzen ist. Dass die kompensierten Reaktionen ausschlaggebend sind, leuchtet ein, weil es um die Minimierung der Zusatzbelastung bei gegebenem Aufkommen geht, die durch den Substitutionseffekt bestimmt ist. Es gibt für (B.I.4) eine (nicht allzu) intuitive Beschreibung von Samuelson[11]: Optimale Steuern liegen vor, wenn, ausgehend von diesen Steuern, eine (kleine) proportionale Erhöhung aller Steuern die kompensierte Nachfrage nach allen Gütern um einen einheitlichen Prozentsatz reduzieren würde.

Wichtiger als diese Intuition sind einige spezifische Aussagen:

(i) Im Fall, dass in den Nachfragefunktionen alle Kreuzpreiseffekte null sind, ergibt sich die "Inverse-Elastizitäten-Regel", gemäß der jene Güter hohe Steuersätze tragen sollen, deren Preiselastizität gering ist.

---

[10] Eine Herleitung findet sich in Anhang 1. Für Herleitungen und Interpretationen in einem Modell mit mehr als zwei Konsumgütern siehe etwa Atkinson und Stiglitz 1980, Rose, Wenzel und Wiegard 1981, Homburg 2000.
[11] Atkinson und Stiglitz 1980, p. 373.

(ii) Ein einheitlicher Steuersatz auf beide Güter (das entspricht einer Steuer auf Arbeitseinkommen allein, ohne Gütersteuern, vgl. die Überlegungen im Anschluss an (B.I.1)) ist optimal, wenn die Nutzenfunktion des Haushalts schwach separabel zwischen Konsum und Freizeit ist (d. h. die Grenzrate der Substitution zwischen c_1 und c_2 ist unabhängig von F bzw. L) und homothetisch bezüglich der Konsumgüter ist.

(iii) Wenn ein einheitlicher Steuersatz auf beide Güter nicht optimal ist, so soll jenes Gut stärker besteuert werden, das ein Komplementärgut zur Freizeit ist ("Ersatzbesteuerung" der Freizeit, Corlett-Hague 1953).

Ein Problem dieses Modells ist, dass sich im Fall eines repräsentativen Haushalts (bzw. - gleichbedeutend dazu - im Fall identischer Haushalte) nicht begründen lässt, warum der Staat nicht einfach eine Pauschalsteuer (Kopfsteuer) einhebt, es gibt ja kein Verteilungsproblem. Eine mögliche Erweiterung stellt die Betrachtung eines Modells mit mehreren Personen dar, die sich in den Präferenzen und/oder Lohnsätzen unterscheiden. Solange man sich jedoch auf proportionale Steuern beschränkt, ergeben sich keine prinzipiell neuen Einsichten (außer dass Verteilungsaspekte hinzukommen). Im Folgenden wenden wir uns daher der Untersuchung der optimalen (nicht-proportionalen) Einkommensteuer zu.[12]

B.II. Das Mirrlees-Modell

Typischerweise wird in modernen Volkswirtschaften das Einkommen als der wichtigste Indikator für die individuelle Leistungsfähigkeit angesehen und als Grundlage für die Besteuerung herangezogen, wobei der Durchschnitt- und der Grenzsteuersatz sich mit der Höhe des Einkommens ändern (können). Der "richtige" Verlauf der Steuerfunktion stellte seit langem einem wichtigen Diskussionspunkt in der (normativen) Finanzwissenschaft dar, wobei - wie in der Einleitung erwähnt - lange Zeit die Opfertheorie und daraus ableitbare Folgerungen im Zentrum der Überlegungen standen. Mit der Arbeit von Mirrlees (1971) wurde die Bedeutung der Anreizeffekte, die von der Einkommensteuer auf das Arbeitsangebotsverhalten ausgehen, in den Mittelpunkt theoretischer Überlegungen gerückt, während als Verteilungsnorm nicht mehr ein Opferprinzip sondern eine (utilitaristische) soziale Wohlfahrtsfunktion herangezogen wurde. Der Anlass, warum Umverteilung in diesem Modell als wünschbar erscheinen kann, liegt darin, dass sich die Personen in ihrem Lohnsatz, also in ihrer Fähigkeit, Einkommen zu erzielen, unterscheiden.

[12] Erwähnt sei noch ein wichtiges Resultat, das auf Diamond und Mirrlees (1971) zurückgeht: In Modellen mit komplexerer Produktionsstruktur weist ein optimales Steuersystem unter bestimmten Voraussetzungen die Eigenschaft auf, dass es Produktionseffizienz induziert. D. h., die Verzerrung der Preise durch (optimal gewählte) Steuern ändert nichts daran, dass auf dem effizienten Rand der Möglichkeitenmenge produziert wird.

Die ursprüngliche Arbeit von Mirrlees (1971) betrachtete ein Kontinuum von unterschiedlichen Lohnsätzen w, die die Personen charakterisieren. Die Verteilung von w wird dabei durch die Wahrscheinlichkeitsdichte f bzw. die zugehörige Verteilungsfunktion F beschrieben, mit $\underline{w}$, $\overline{w}$ als Unter- und Obergrenze für die Lohnsätze.

Wir betrachten im Folgenden eine sehr einfache Variante des Mirrlees-Modells, nämlich ein Zwei-Personen-Modell,[13] das sich zur Illustration der wichtigsten Eigenschaften recht gut eignet. Dagegen erweitern wir das Modell, indem wir zwei Konsumgüter betrachten. Wir bleiben somit beim im vorigen Abschnitt eingeführten Haushaltsmodell, nehmen aber nun an, dass es zwei Personen gibt, die sich in ihrer Fähigkeit, Arbeit zu leisten und Einkommen zu erzielen, unterscheiden (d. h. in einem Effizienzparameter). Dies drückt sich direkt in unterschiedlichen Lohnsätzen $w^1 < w^2$ aus (d. h. der Lohnsatz pro Effizienzeinheit der geleisteten Arbeit wird mit eins normiert). [14]

Die Nutzenfunktion $U(c_1,c_2,L)$ ist identisch für beide Personen, sie wird als konkav vorausgesetzt. $z^h \equiv w^h L^h$, h = 1, 2 bezeichnet das Bruttoeinkommen (= angebotene Arbeit in Effizienzeinheiten) der beiden Individuen.

Wir benennen mit x^h die Ausgaben (ein *composite commodity* im Sinn von Hicks) für die beiden Konsumgüter, bei vorgegebenen Konsumentenpreisen q_1, q_2 $(q_i = p_i + \hat{\tau}_i)$ und definieren die (direkt-indirekte) Nutzenfunktion $v^h(z,x;q_1,q_2) \equiv \max\left\{U(c_1,c_2,z/w^h)/\,q_1c_1 + q_2c_2 \leq x\right\}$. In dieser Formulierung kann x als Nettoeinkommen interpretiert werden; v^h beschreibt den erreichbaren Nutzen für Person h, wenn sie ein Nettoeinkommen x bekommt (und für den Konsum der beiden Güter zu Preisen q_1, q_2 verwendet) und gleichzeitig ein bestimmtes Bruttoeinkommen z erarbeiten muss, also dafür die Arbeitszeit z/w^h aufwenden muss (die partielle Ableitung $\partial v^h/\partial z$ ist daher negativ: für höheres Bruttoeinkommen muss mehr Arbeitszweit aufgewendet werden; $\partial v^h/\partial x$ ist positiv).

Ein wichtiger Punkt, der in diesem Modell sehr klar präzisiert werden kann, betrifft die Frage, ob das (Brutto-)Einkommen z^h einen geeigneten Indikator für die Leistungsfähigkeit w^h darstellt. Nur in diesem Fall entspricht ja eine Einkommensteuer, bei der die Steuerschuld mit dem

[13] Für eine Darstellung des kontinuierlichen Modells siehe u. a. auch Ebert (1992), Brunner (1990), Boadway u. a. 2000. In Homburg (2001) wird der Zusammenhang zwischen dem diskreten Modell (mit endlich vielen Fähigkeitsniveaus) und dem kontinuierlichen erörtert.
[14] Im Folgenden bezeichnet der obere Index die Personen.

Bruttoeinkommen steigt, der gewünschten Verteilungszielrichtung von höherer zu niedrigerer Leistungsfähigkeit. Die entscheidende Bedingung an die Nutzenfunktion U (und daraus hergeleitet an v^1, v^2), dass dies zutrifft lautet[15]

$$\text{AM:}\quad MRS^{1}_{zx} > MRS^{2}_{zx},\ \text{für jedes } z,x,q_1,q_2, \tag{B.II.1}$$

wobei $MRS^{h}_{zx} \equiv -(\partial v^h/\partial z)/(\partial v^h/\partial x)$ die Grenzrate der Substitution zwischen Brutto- und Nettoeinkommen bezüglich v^h darstellt; sie muss für die weniger produktive Person größer als für die leistungsfähigere Person sein, und zwar bei jeder beliebigen $(\bar{x},\bar{z})$-Kombination (und gegebenen Preisen q_1, q_2). In Abbildung 1 ist dies grafisch dargestellt:

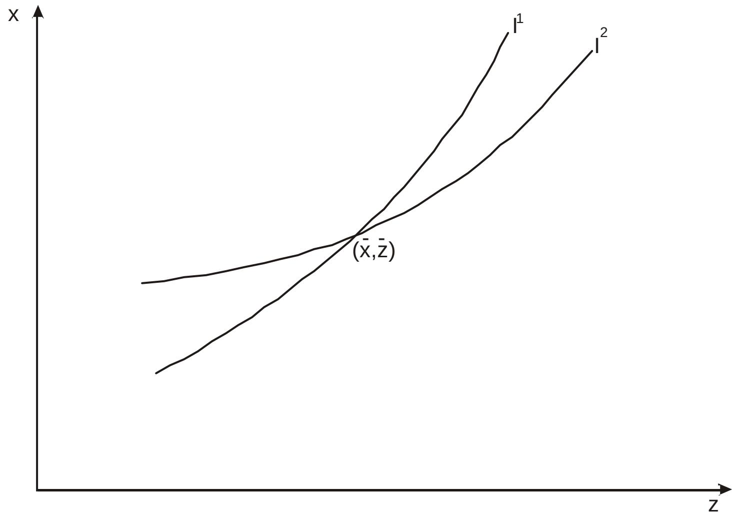

Abbildung 1

I^1 und I^2 bezeichnen je eine Indifferenzkurve von Person eins bzw. Person zwei. AM besagt, dass im Schnittpunkt $(\bar{z},\bar{x})$ zweier solcher Kurven stets die zur Person eins gehörende Indifferenzkurve steiler verläuft als die zur Person zwei gehörende.[16]

Dass die Präferenzen (Nutzenfunktionen) der Individuen die Bedingung AM erfüllen ist intuitiv plausibel: um eine Einheit mehr Bruttoeinkommen zu erzielen, benötigt Person zwei weniger zusätzliche Arbeitszeit als Person eins, daher braucht sie zur Kompensation (um auf gleichem Nutzenniveau zu bleiben) vermutlich weniger zusätzliches Nettoeinkommen. AM impliziert, dass bei jeder Einkommensteuerfunktion die Person mit höherer Fähigkeit eine solche Arbeitszeit wählt,

[15] AM steht für *agent monotonicity* (Seade 1982).
[16] Daher können sich I^1 und I^2 auch nur einmal schneiden, AM ist somit eine *single-crossing* condition.

dass sie ein mindestens so großes Bruttoeinkommen erzielt wie eine Person mit niedrigerer Fähigkeit. [17]

Bei den folgenden Überlegungen konzentrieren wir uns zunächst auf die Festlegung der Einkommensteuer und nehmen dazu die beiden Gütersteuern $\hat{\tau}_1, \hat{\tau}_2$ als fix gegeben an, zum Beispiel mit dem Wert null. Die Zielsetzung des Staates (eines wohlwollenden Planers) ist, wie erwähnt, die Maximierung einer (utilitaristischen) sozialen Wohlfahrtsfunktion, wobei er ein gegebenes Steueraufkommen g erzielen möchte. Was sind seine Instrumente? Eine erste Möglichkeit wäre, eine bestimmte funktionale Form für die zu bestimmende Einkommensteuer vorzugeben und dann jene Parameterwerte für diese Funktion zu suchen, so dass die Zielfunktion maximiert wird. Dem entspricht das Vorgehen für die Bestimmung der optimalen linearen Einkommensteuer. Diese ist durch zwei Parameter charakterisiert: einen fixen Steuerbetrag α und einen marginalen Steuersatz β: $x = -\alpha + (1-\beta)z$.[18]

In Abbildung 2 ist ein möglicher Verlauf einer linearen Einkommensteuer im Brutto-Netto-Einkommensdiagramm skizziert, wobei $\alpha < 0$ (garantiertes Mindesteinkommen) und $0 < \beta < 1$ unterstellt ist. Die Paare (z^h, x^h) bezeichnen die Entscheidungen der Individuen bei dieser Steuerfunktion; der jeweilige vertikale Abstand zur 45°-Linie (Brutto- = Nettoeinkommen) die Höhe der gezahlten Steuer (negativ für Person eins, positiv für Person zwei).

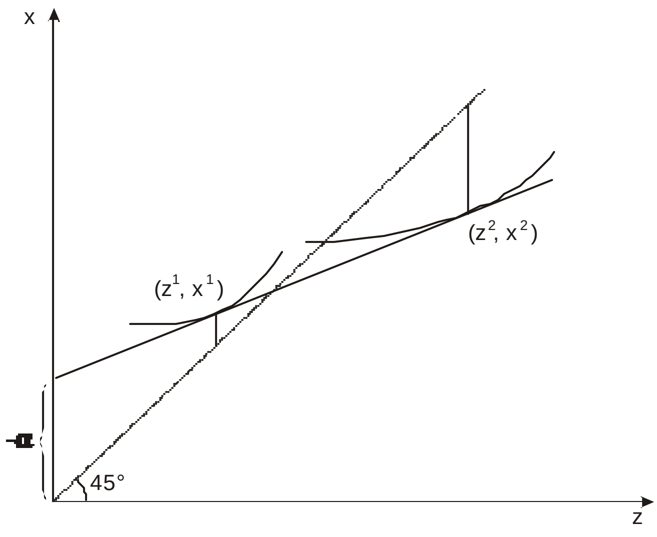

Abbildung 2

Die Idee von Mirrlees ist jedoch, die bestmögliche Steuerfunktion ohne jede funktionale Vorgabe zu bestimmen. Die Überlegung lautet: Um den Wert der Zielfunktion (Nutzensumme) zu kennen,

[17] Für eine ausführliche Diskussion siehe Brunner 1989, p 26ff.
[18] Zur optimalen Festlegung von α und β siehe Sheshinski (1972), Hellwig (1986), Brunner (1989).

muss der Staat bei jeder möglichen Steuerfunktion die Brutto- und Nettoeinkommenspositionen berechnen, die die Personen bei dieser Steuerfunktion wählen.[19] Dann liegt der umgekehrte Zugang nahe: Suche gemäß der Zielfunktion direkt die besten Brutto- und Nettoeinkommenspositionen der Haushalte, unter der Bedingung, dass es dazu eine Steuerfunktion gibt, bei der die Personen eben diese Brutto- und Nettoeinkommenspositionen selbst wählen.

Nun kann man sich leicht überlegen, dass in unserem einfachen Modell zwei Paare (z^1,x^1), (z^2x^2) die letztere Bedingung (Existenz einer Steuerfunktion) erfüllen, wenn Folgendes gilt:

$$v^1(z^1,x^1) \geq v^1(z^2,x^2), \tag{B.II.2}$$

$$v^2(z^2,x^2) \geq v^2(z^1,x^1). \tag{B.II.3}$$

Das heißt, wenn Person eins beim Paar (z^1,x^1) mindestens so gut gestellt ist wie beim Paar (z^2,x^2), und Analoges für Person zwei gilt, dann existiert eine Steuerfunktion σ, mit $x = z - \sigma(z)$, sodass Person eins selbst das Paar (z^1,x^1) und Person zwei selbst das Paar (z^2,x^2) wählt. (B.II.2) und (B.II.3) sind die in der Literatur als *self-selection constraints* bezeichneten Bedingungen (Stiglitz 1982).[20]

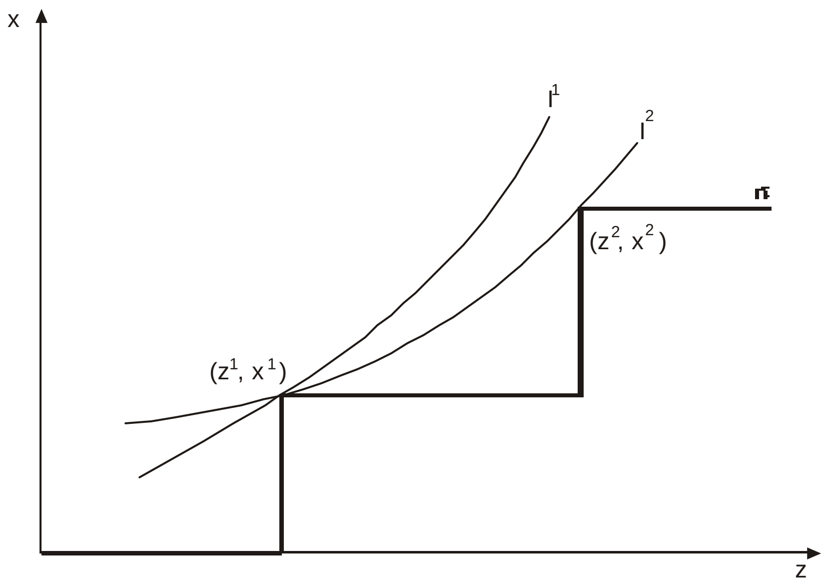

Abbildung 3

[19] Ähnlich wie im Ramsey-Modell: auch dort kann der Staat die optimalen Steuersätze nur finden, indem er im Prinzip für alle möglichen Steuersätze die zugehörigen vom repräsentativen Haushalt selbst gewählten Güterbündel bestimmt (d. h., er kennt die Nachfragefunktion und die indirekte Nutzenfunktion).

[20] In einem Modell mit n Personen lauten sie analog: $v^h(z^h,x^h) \geq v^j(z^jx^j)$, für alle h, j = 1, ..., n und in der gleichen Weise lassen sie sich auch bei einem Kontinuum von Personen definieren.

Abbildung 3 zeigt eine Situation, bei der (B.II.3) strikt (mit Gleichheit) erfüllt ist, ((z^1,x^1) und (z^2,x^2) liegen auf der gleichen Indifferenzkurve I^2 von Person zwei), (B.II.2) dagegen nicht strikt (beachte, dass wegen AM der Anstieg von I^1 in (z^1,x^1) größer ist als der von I^2).

Als eine mögliche Steuerfunktion ist eine Stufenfunktion σ eingezeichnet:

$$\sigma(z) = \begin{cases} z & \text{für } x < z^1, \\ z - x^1 & \text{für } z^1 \leq z < z^2, \\ z - x^2 & \text{für } z^2 \leq z, \end{cases} \qquad \text{bzw. } x = z - \sigma(z) = \begin{cases} 0 & \text{für } x < z^1, \\ x^1 & \text{für } z^1 \leq z < z^2, \\ x^2 & \text{für } z^2 \leq z. \end{cases}$$

Die Beachtung der Selbstselektionsrestriktionen bringt die Informationsbeschränkung des Staates zum Ausdruck: Er kann nur das Bruttoeinkommen, nicht aber die Lohnsätze selbst beobachten und die Steuer daran anknüpfen. Somit gelangt er zu einer *second-best*-Lösung, während die *first-best*-Lösung direkt an den Lohnsätzen anknüpfen würde - eine nach Lohnsätzen differenzierte Pauschalsteuer (beachte, dass die Lohnsätze bzw. Fähigkeiten in diesem Modell als fix gegeben angenommen sind).[21]

Schließlich bleibt noch die Ressourcenbeschränkung zu formulieren, die der Planer zu beachten hat:

$$x^1 + x^2 \leq z^1 + z^2 - (g - \hat{\tau}_1(c_1^1 + c_1^2) + \hat{\tau}_2(c_2^1 + c_2^2)) \,. \tag{B.II.4}$$

Die Summe der Nettoeinkommen darf nicht größer als die Summe der Bruttoeinkommen (= Arbeitsleistung in Effizienzeinheiten) minus staatlichem Ressourcenanspruch (vermindert um das Gütersteueraufkommen) sein.

Wenn wir mit ρ^1 und ρ^2 die Gewichte der beiden Personen in der utilitaristischen Zielfunktion bezeichnen, so können wir das Optimierungsproblem des Staates nun formulieren:

$$\max_{z^1,z^2,x^1,x^2} \rho^1 v^1(z^1,x^1) + \rho^2 v^2(z^2,x^2) \tag{B.II.5}$$

s. t. (B.II.2), (B.II.3), (B.II.4)

$$z^h, x^h \geq 0,\ h = 1, 2. \tag{B.II.6}$$

[21] Unterstellt ist im Mirrlees-Modell allerdings, dass der Staat die Verteilung der Lohnsätze kennt.

B.III. Ergebnisse, Erweiterungen und Kritik des Mirrlees-Modells

B.III.1. Eigenschaften der optimalen Einkommensteuerfunktion

Welche Aussagen lassen sich aus diesem Modell für den optimalen Steuerverlauf herleiten? Entscheidend dafür ist die Frage, welche der beiden Selbstselektionsbedingungen (B.II.2), (B.II.3) bei der optimalen Lösung *bindend* ist, das heißt, mit Gleichheit erfüllt. (Wie man sich leicht überlegt, können wegen der single-crossing condition nicht beide gleichzeitig erfüllt sein.[22]) Die Antwort darauf hängt mit der Verteilungszielsetzung des Staates zusammen. Wenn er von oben nach unten umverteilen möchte, so ist offensichtlich (B.II.3) die relevante Bedingung: Der Umverteilungsvorgang (Verringerung des Nettoeinkommens für Person eins und Erhöhung für Person zwei) ist dadurch begrenzt, dass es schließlich Person zwei für besser finden könnte, das der Person eins zugedachte (z^1,x^1)-Bündel zu wählen, an Stelle des ihr selbst zugedachten (ein solches Verhalten wird in der Literatur manchmal als "mimicking" bezeichnet; man beachte, dass Person zwei weniger lang zu arbeiten braucht, um z^1 zu erwirtschaften, als Person eins).

Man kann zeigen, dass Normalität von Freizeit und strikte Konkavität von U zusammen mit $\rho^1 \geq \rho^2$ (oder Konkavität und $\rho^1 > \rho^2$) und eine hinreichende Bedingung darstellen, dass die utilitaristische soziale Wohlfahrtsfunktion tatsächlich eine Umverteilung in der beschriebenen Weise impliziert (Brunner 1995). In Abbildung 4 ist eine solche Lösung skizziert: Person zwei ist zwischen den Bündeln (z^1,x^1) und (z^2,x^2) indifferent.[23] Als charakterisierende Bedingungen für die optimalen Bündel (z^1,x^1), (z^2,x^2)[24] ergeben sich in unserem Zwei-Personen-Modell (wobei $\hat{\tau}_1 = \hat{\tau}_2 = 0$ vorausgesetzt sei):

$$MRS^1_{zx}(z^1, x^1) < 1, \qquad \text{(B.III.1)}$$

$$MRS^2_{zx}(z^2, x^2) = 1 \qquad \text{(B.III.2)}$$

(B.III.1) besagt, dass im Optimum die Person mit dem geringeren Lohnsatz eine Position einnimmt, bei der die Grenzrate der Substitution zwischen Brutto- und Nettoeinkommen kleiner als eins ist. Das kann man in der Weise interpretieren, dass sie sich einem positiven Grenzsteuersatz gegenübersieht. (Bei einer gegebenen, differenzierbaren Steuerfunktion σ wird gemäß der üblichen Grenzratenbedingung die Arbeits-Freizeit-Entscheidung einer Person h im z-x-Diagramm

[22] Außer im Sonderfall, dass beide Personen das gleiche (z,x)-Bündel bekommen. Dies kann im Zwei-Personen-Modell nicht optimal sein, wie in Anhang 2 gezeigt wird.
[23] Dabei ist - wie in solchen Modellen generell üblich - unterstellt, dass Person zwei bei Indifferenz tatsächlich das ihr zugedachte Bündel wählt.
[24] Zur Herleitung siehe Anhang 2 sowie, u. a., Guesnerie und Seade (1982), Brunner (1989), Homburg (2000).

zu einer Brutto-Netto-Einkommensposition führen, bei der $MRS^h_{zx}(z^h, x^h) = 1 - \sigma'(z^h)$ gilt, wie dies in Abbildung 2 für den Fall einer linearen Steuerfunktion illustriert ist. $MRS^h_{zx} < 1$ ist daher äquivalent zu $\sigma'(z^h) > 0$, und dies motiviert die obigen Interpretation.)

(B.III.3) beschreibt das berühmte und umstrittene Resultat, dass für die Person mit dem höchsten Lohnsatz der Marginalsteuersteuersatz im Optimum null ist (Sadka 1976, Seade 1977).

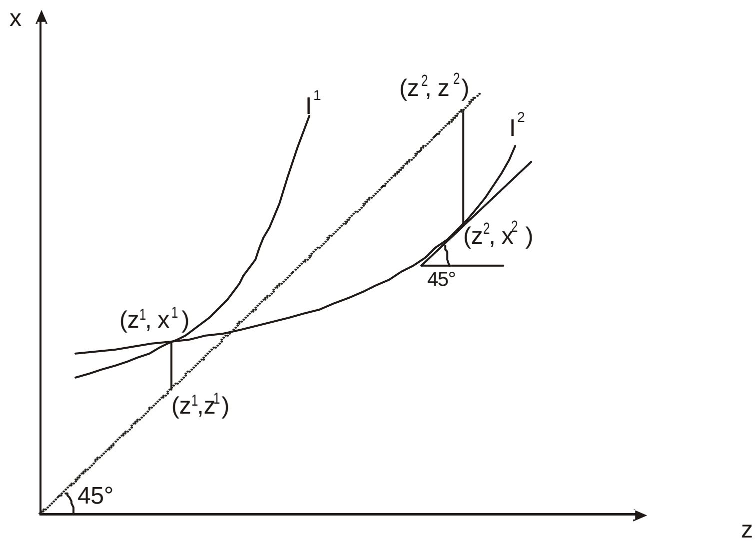

Abbildung 4

In Abbildung 4 ersieht man (B.III.3) daran, dass v^2 in (z^2,x^2) einen Anstieg von 1 aufweist, während v^1 in (z^1,x^1) mit weniger als 1 ansteigt (B.III.1). Die Differenz zwischen (z^2,x^2) und (z^2,z^2) stellt die entrichtete Steuer von Person zwei dar, jene zwischen (z^1,x^1) und (z^1,z^1) die Steuer von Person eins (im gezeichneten Fall ist das ein Transfer an Person eins, der Rest geht als Aufkommen an den Staat). Erwähnt sei noch, dass auch eine Lösung $z^1 = 0$ optimal sein kann, die also auf dem Rand liegt: Person eins arbeitet nicht.

Auch in einem allgemeineren Modell mit einer größeren Zahl von Personen erhält man aus der Theorie allein keine wesentlich weiter gehenden Aussagen.[25] Daher haben verschiedene Autoren versucht, durch Simulationsrechnungen genauere Eigenschaften des optimalen Einkommensteuertarifs herzuleiten (u. a. Tuomala 1990). Erste Rechnungen dazu finden sich bereits in der Originalarbeit von Mirrlees (1971), für das Modell mit kontinuierlicher Verteilung der Fähigkeiten. Offensichtlich sind die Ergebnisse sensibel hinsichtlich der unterstellten Nutzenfunktion (Arbeitsangebotsreaktion), der Intensität der Umverteilungszielsetzung sowie der

[25] Von theoretischem Interesse ist allerdings, dass bei mehr Personen die optimale Lösung die Eigenschaft haben kann, dass mehrere Personen das gleiche (z,x)-Paar zugedacht bekommen (*bunching*). Siehe etwa Ebert (1992).

Fähigkeitsverteilung; insgesamt ergibt sich bei Mirrlees ein Verlauf des Marginalsteuersatzes gemäß einem invertierten U oder sogar ein nahezu konstanter Marginalsatz als plausibles Ergebnis. (Abbildung 5 zeigt den sich ergebenden Zusammenhang von Brutto- und Nettoeinkommen für einige Parameterwerte, wobei eine Lognormalverteilung der Fähigkeiten unterstellt ist.)

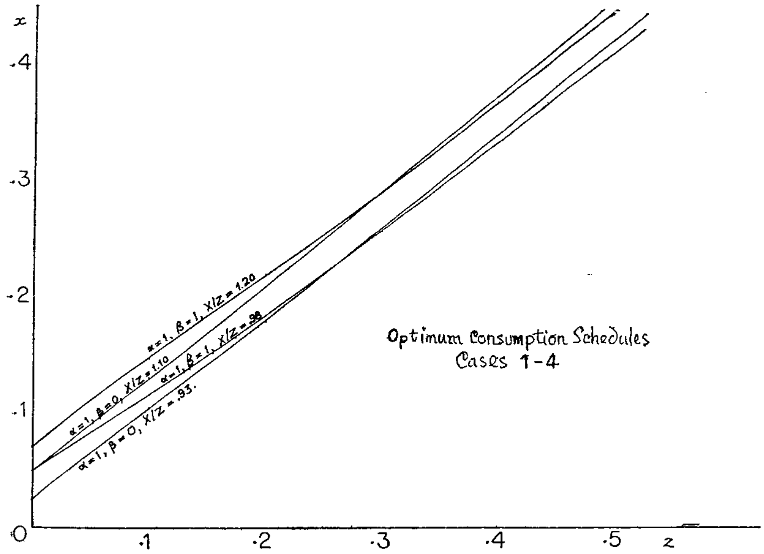

FIGURE 5

Abbildung 5

(Quelle: Mirrlees 1971)

Eine Intuition, warum der optimale Grenzsteuersatz gegen das obere Ende sinken dürfte, liefert folgende Überlegung in einem Modell mit vielen Personen: Die Erhöhung des Marginalsatzes an einer bestimmten Stelle der Steuerfunktion, d. h. bei einer bestimmten Person (deren Lohnsatz mit $\tilde{w}$ bezeichnet sei) bringt zwei Effekte mit sich, nämlich (i) eine (unerwünschte) stärkere Verzerrung der Arbeitsangebotsentscheidung dieser Person, (ii) ein höheres Steueraufkommen (ein höheres Durchschnittsteuersatz) von allen Personen oberhalb $\tilde{w}$, ohne dass deren Arbeitsangebotsentscheidung dadurch zusätzlich verzerrt würde. Der zweite (und für Umverteilungszwecke erwünschte) Effekt wird immer geringer, je größer $\tilde{w}$ ist, was schließlich erklärt, warum sich in den Simulationen ein Abnehmen des Marginalsatzes gegen das obere Ende hin ergibt.[26]

Bei Annahme einer nach oben offenen Verteilung der Fähigkeiten existiert kein höchstes Fähigkeitsniveau, für das das Resultat des verschwindenden Grenzsteuersatzes exakt gelten

[26] In einem Modell mit kontinuierlicher Lohnsatzverteilung lässt sich die quantitative Bedeutung von (i) durch $f(\tilde{w})$, jene von (ii) durch $1 - F(\tilde{w})$ beschreiben. Somit ist ein höherer Grenzsteuersatz umso eher erwünscht, je größer der Wert von (1-F(w))/f(w). Er wird null bei $\overline{w}$, falls $f(\overline{w}) > 0$.

würde. Es zeigt sich, dass für den Verlauf des optimalen Grenzsteuersatzes im oberen Einkommensbereicht die genaue Gestalt der Fähigkeitsverteilung maßgeblich ist. Dieser Zusammenhang wurde in jüngerer Zeit von Diamond (1998) analysiert, der die Diskussion um den Verlauf des optimalen Einkommensteuertarifs wieder aufgenommen hat.[27] Er zeigte, dass sich bei einer quasilinearen Nutzenfunktion (linear im Nettoeinkommen) unter der Annahme einer konstanten Arbeitsangebotselastizität bei einer Pareto-Verteilung der Fähigkeiten ein auch am oberen Ende steigender Marginalsatz als optimal ergibt. Allerdings hängt dieses Resultat wesentlich am konstanten Grenznutzen des Einkommens, worauf Daran und Strawczynski (2000) hingewiesen haben. (Abbildung 6)

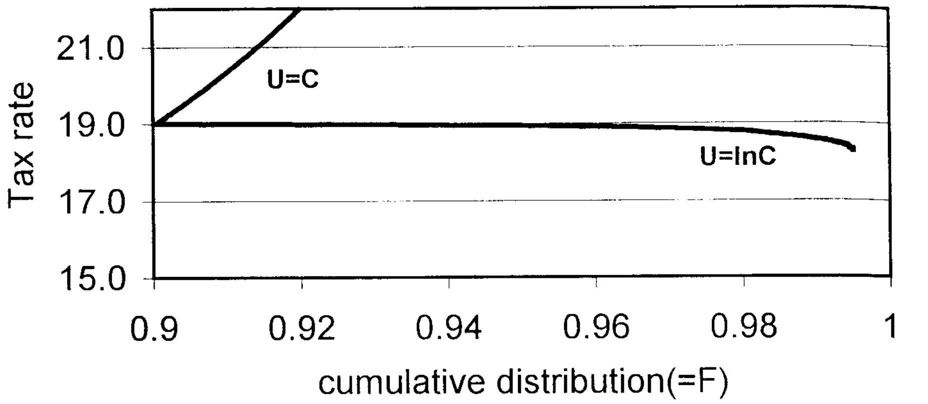

FIGURE 2A. OPTIMAL TAXES AT HIGH LEVELS OF SKILL
THE CASE OF PARETO DISTRIBUTION
(ALFA = 1.5 AS IN DIAMOND, 1998)

Abbildung 6

(Quelle: Daran und Strawczynski 2000)

In einer neueren Arbeit hat Saez (2001) das empirische beobachtbare Einkommen an Stelle des unbeobachtbaren Fähigkeitsniveaus als zentrale Größe gewählt und Formeln entwickelt, wie der Verlauf des optimalen nichtlinearen Steuertarifs durch die (kompensierten und nichtkompensierten) Elastizitäten des zu versteuernden Einkommens in Bezug auf den jeweiligen Grenzsteuersatz charakterisiert werden kann.[28] Ein besonders einfacher Ausdruck für den optimalen Grenzsteuersatz im oberen Einkommensbereich ergibt sich, wenn man annimmt, dass er ab einem gewissen Niveau konstant bleibt. Unter Verwendung einer Paretoverteilung der Einkommen im oberen Ende (die die tatsächliche Verteilung der Arbeitseinkommen in den USA sehr gut annähert) erhielt Saez, je nach Annahme über die Höhe der Elastizitäten und der gewünschten

[27] Siehe auch Saez (2001).

[28] Die Höhe der (unkompensierten) Elastizitäten des steuerlichen Einkommens in Bezug auf den Grenzsteuersatz wurde in einer Reihe von Arbeiten empirisch untersucht. Während Feldstein (1995) einen Wert von über eins ermittelte, fanden spätere Untersuchungen deutlich geringere Werte (zwischen null und 0,5), vor allem auf längere Frist (Goolsbee 2000, Gruber und Saez 2002).

Umverteilungszielsetzung, relativ hohe Grenzsteuersätze für den oberen Bereich (zwischen 30% und 90%).

Die Ergebnisse der Simulationen von Saez zum gesamten Verlauf des optimalen Grenzsteuersatzes für das ursprüngliche Mirrlees-Modell sind in Abbildung 7 wiedergegeben. Dabei ist die Fähigkeitsverteilung so kalibriert, dass sich die tatsächliche Verteilung der Bruttoeinkommen in den USA ergibt (Paretoverteilung).[29] Es ergeben sich vergleichsweise hohe Marginalsätze. Die beiden Diagramme unterscheiden sich bezüglich des Einkommenseffekts, der im linken Diagramm mit null angenommen wird. Es sind jeweils die optimalen Verläufe für Werte der kompensierten Elastizität von 0,25 bzw. 0,5 eingezeichnet. Die strichlierten Werte zeigen den optimalen Marginalsatz bei einem linearen Tarif.

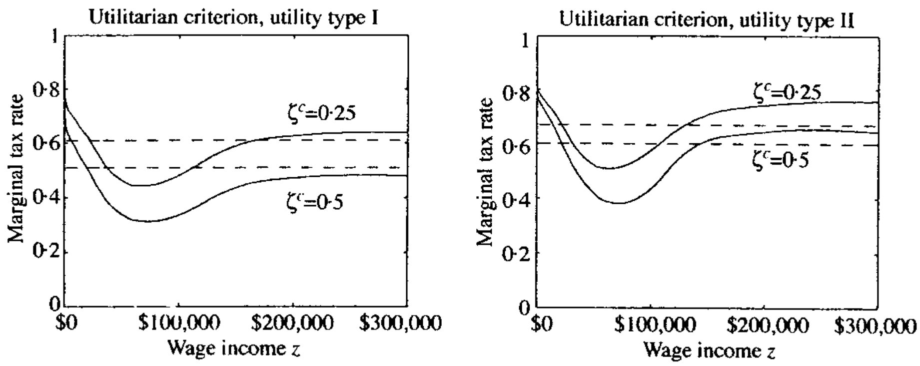

Abbildung 7
(Quelle: Saez 2002)

B.III.2. Das untere Ende des Steuertarifs

Das Mirrlees-Model beinhaltet, wie in Abbildung 4 ersichtlich, die Möglichkeit negativer Steuern. Es stellt also eigentlich einen Ansatz für die Suche nach einem integrierten Steuer- und Sozialsystem dar, in einem Modellrahmen, bei dem die Arbeitszeitentscheidung im Mittelpunkt des Interesses steht. Offensichtlich treten negative Steuern, wenn überhaupt am unteren Ende der Einkommensbezieher auf und es liegt daher nahe, den Steuerverlauf in diesem Bereich besonders zu studieren.

[29] Die von Mirrlees in seinen Simulationen verwendete Lognormalverteilung ist eine der Ursachen für sein Ergebnis abnehmender (und niedriger) optimaler Grenzsteuersätze im oberen Bereich.

Als theoretisches Ergebnis relevant ist hier die Aussage, dass der optimale Marginalsteuersatz nie negativ sein kann bzw. sogar mit Ausnahme des obersten Niveaus immer strikt positiv ist.[30] Statt Personen mit niedrigem Lohnsatz durch Lohnsubventionierungen einen Anreiz zum Arbeiten zu geben, ist es im Sinn der sozialen Wohlfahrtsfunktion effizienter, ihnen direkt eine Zahlung zukommen zu lassen. Simulationsrechnungen zeigen tendenziell relativ hohe Marginalsätze im unteren Bereich. Dieses Resultat entspricht dem Problem, das mit Transferzahlungen an Personen mit niedrigem oder gar keinem Einkommen stets verbunden ist: Die Transferzahlungen müssen mit steigenden Einkommen abgebaut werden und das erfordert hohe Marginalsätze am unteren Ende (Armutsfalle).

In einer jüngeren Arbeit hat Saez (2002a) ein Modell der optimalen Einkommensteuer vorgestellt, das als Lösung einen Verlauf mit negativen Marginalsätzen am unteren Ende liefert, wie es für Programme zur Subventionierung der Arbeit[31] typisch ist. Das Modell beruht auf einem älteren Ansatz von Diamond (1980), bei dem der Aspekt betont wird, dass viele Arbeitnehmer kaum eine freie Wahl der Arbeitszeit vornehmen können. Dementsprechend gibt es in diesem Modell eine Menge von Arbeitsplätzen, die unterschiedliches, aber fixes Bruttoeinkommen bieten. Durch die Einkommensteuer (bzw. den Transfer) wird das verfügbare Einkommen beeinflusst. Die Personen unterscheiden sich in ihren Präferenzen - manche benötigen ein höheres Nettoeinkommen als andere, um einen Arbeitsplatz mit einem bestimmten Bruttoeinkommen (= Arbeitsanstrengung) anzunehmen. Saez bezeichnet als *extensive* Arbeitsangebotsreaktion die prinzipielle Entscheidung Entscheidung einer Person, einen Arbeitsplatz zu akzeptieren und nimmt an, dass sie nur von der Differenz des verfügbaren Einkommens bei diesem Arbeitsplatz und dem garantierten Mindesteinkommen ohne Arbeit abhängt. Die Zielsetzung des Staates ist die Maximierung einer gewichteten sozialen Wohlfahrtsfunktion (wobei unterstellt wird, dass die Gewichte umso kleiner werden, je höher das Einkommen ist). In diesem Rahmen kommt Saez zum Resultat, dass die optimale Lösung am unteren Ende einen negativen Marginalsteuersatzes aufweist. Dies ist etwa in Abbildung 8 ersichtlich, wo für einen Parameterwert $\eta = 1$ die Netto-Bruttoeinkommenskurve steiler als 45 ° verläuft.[32]

Wenn auch die Konstruktion des Modells etwas ungewöhnlich erscheint, so ist doch anzuerkennen, dass es das empirisch gesicherte Resultat einer größeren Elastizität der prinzipiellen Arbeitsangebotsentscheidung (im Vergleich zur Arbeitszeitentscheidung) bezüglich

[30] Allerdings tritt im Modell mit kontinuierlicher Fähigkeitsverteilung noch der Fall auf, dass der optimale Marginalsatz auch beim untersten Niveau null ist, sofern alle Personen im Optimum eine positive Arbeitszeit wählen (Seade 1977, siehe auch Brunner 1990, Homburg 2001).
[31] Etwa dem Earned Income Tax Credit (EITC) in den USA.
[32] Saez modelliert auch den Fall der *intensiven* Arbeitsangebotsreaktion, bei der die Wahl des Arbeitsplatzes von der Nettolohndifferenz zu anderen Arbeitsplätzen (mit höherem oder geringerem Bruttoeinkommen) abhängt. In diesem Fall erweist sich eine negative Einkommensteuer als unteren Ende als optimale Lösung.

des Lohnsatzes ernst nimmt. [33] Es liefert eine interessante wohlfahrtstheoretische Begründung für Subventionierungen der Arbeit, die sich im Mirrlees-Modell nicht finden lässt.[34,35]

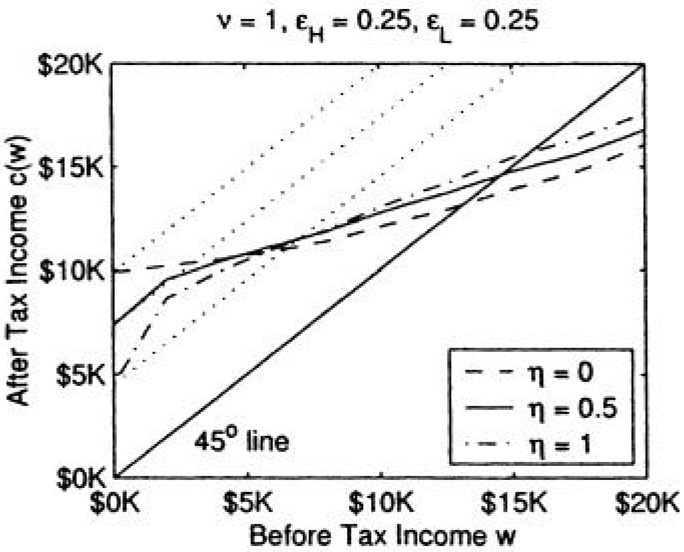

Abbildung 8

(Quelle: Saez 2002)

B.III.3. Kritik des Mirrlees-Modells

Wie bei jedem theoretischen (oder auch empirischen) Resultat, das sich als Konsequenz einer bestimmten Modellformulierung ergibt, ist seine Bedeutung von der Plausibilität der Annahmen abhängig. Im Folgenden werden einige kritische Punkte angesprochen.

Eine Einsicht aus dem Modell, die man bedauern mag, aber auch nicht überraschen sollte, ist das geringe Ausmaß an allgemeinen Erkenntnissen, die es liefert. Sie bestehen im Wesentlichen aus der durchgehenden Nichtnegativität des Marginalsatzes und dessen Verschwinden für die Person mit dem höchsten Fähigkeitsniveau. Für genauere Aussagen über den Steuerverlauf muss man auf Simulationsrechnungen zurückgreifen, die offensichtlich von der jeweiligen Parameterwahl abhängige Ergebnisse liefern. Da kaum jemand erwarten wird, dass man durch empirische Untersuchungen tatsächlich zu einer Kenntnis der individuellen Nutzenfunktion kommen kann, noch dass man etwas Konkretes zur Form der sozialen Wohlfahrtsfunktion sagen kann, bleiben genauere Aussagen mit einer gewissen Willkür behaftet. Man kann diesen Sachverhalt aber auch

[33] In einer Studie zur Reaktion der Haushalte auf Änderungen des marginalen Steuersatzes finden Gruber und Saez (2002) einen relativ geringen Effekt des hohen impliziten Grenzsteuersatzes in jenem Einkommensbereich, wo der EITC ausläuft.
[34] Zu einer Kritik siehe Homburg 2002.
[35] Eine andere Abweichung vom Mirrlees-Modell, nämlich die Hinzunahme von Armut als "public bad" in die staatliche Zielfunktion liefert ebenfalls eine Begründung für einen negativen Marginalsatz am unteren Ende (Wane 2001).

aus einer anderen Perspektive betrachten: die genannten allgemeinen Aussagen ergeben sich allein aus einer schwachen Umverteilungsforderung (von oben nach unten, impliziert die Nichtnegativität des Marginalsatzes), kombiniert mit der Forderung nach Pareto-Effizienz (impliziert den Null-Marginalsatz am oberen Ende, siehe Seade 1977, Sadka 1976). Alle darüber hinausgehenden Eigenschaften des Steuertarifs setzen eine genauere Festlegung der normativen Vorgaben (einschließlich der Fixierung der individuellen Nutzenfunktion) voraus.

Natürlich kann man die allgemeinen Resultate für inakzeptabel halten, weil man eine Subventionierung des Arbeitseinkommens für Personen mit geringem Lohnsatz sowie eine direkte Progression des Steuertarifs für richtig erachtet, ohne sich um deren theoretische Optimalität zu kümmern. Immerhin bietet die Auseinandersetzung mit dem Modell und seinen Resultaten die Möglichkeit, genau herauszuarbeiten, welche unterschiedlichen Annahmen zu dieser Beurteilung führen.

Eine zentrale Frage ist dabei offensichtlich, in wie weit das ökonomische Standardmodell der Arbeits-Freizeit-Entscheidung das tatsächliche Arbeitsangebotsverhalten adäquat beschreibt. Dazu lassen sich natürlich eine Reihe von Einwänden vorbringen, etwa, dass für viele Personen (v. a. am oberen Ende der Einkommen) die materielle Entlohnung eine untergeordnete Motivation für ihre Arbeitszeit (und -intensität) spielt, während intrinsische Faktoren wichtiger sind.[36] Berichte über stark verminderte Lebensqualität durch Arbeitslosigkeit (unabhängig vom Einkommensausfall) weisen ebenfalls in diese Richtung.

Modelle, die ein anderes Entscheidungsverhalten der Personen unterstellen, haben tatsächlich zu anderen als den oben beschriebenen Resultaten für den Verlauf der optimalen Einkommensteuer geführt. So erweist sich eine progressive Einkommensteuer als Pareto-effizient, wenn die Personen ihr relatives Einkommen (im Vergleich zu anderen) als wesentlich ansehen oder Status-Effekte eine Rolle spielen.[37] Auch eine andere Entscheidungssituation, nämlich die Einbeziehung von Unsicherheit bezüglich der Höhe zukünftiger Einkommen kann zu einer progressiven Einkommensteuer als optimaler Lösung führen, wegen des damit verbundenen Versicherungseffekts (Varian 1980, Eaton and Rosen, Strawczynski 1998).[38]

Wie oben erwähnt, lassen sich im Modell der optimalen nichtlinearen Einkommensteuer aus dem Pareto-Kriterium allein wenig Aussagen über den optimalen Steuertarif herleiten. Das konkrete

[36] Jedenfalls dürften die Individuen heterogen sein und unterschiedliche Präferenzen bezüglich Einkommen und Freizeit haben. Siehe dazu Sandmo 1993.
[37] Corneo (2002), Ireland (2001). Zu ähnlichen Modellen siehe Boskin und Sheshinksy (1978) oder Oswald (1983).
[38] Weitere in der Literatur behandelte Modellvarianten betreffen die Betrachtung nicht substitutiver Arbeitstypen sowie Komplementarität des Arbeitseinsatzes mit Kapital (Stiglitz 1982, Feldstein 1973).

Ergebnis von Mirrlees, nämlich der nahezu konstante Grenzsteuersatz (der nicht unplausibel erscheint und daher eine gewisse Akzeptanz fand), beruht auf der Verwendung einer konkreten utilitaristischen sozialen Wohlfahrtsfunktion, die unter Berücksichtigung der self-selection-constraints und der Ressourcenbeschränkung maximiert wird. Die self-selection-constraints ergeben sich aufgrund der in B.II. erörterten Informationsasymmetrie (wie üblich in principle-agent-Modellen): jede Person kennt ihren eigenen Lohnsatz, aber der Staat kann ihn nicht direkt beobachten, sondern nur das Einkommen (das Produkt von Lohnsatz und Arbeitszeit), daher kann er nur eine second-best-Lösung implementieren. Die first-best-Lösung wäre - wie erwähnt - eine nach Fähigkeiten differenzierte Pauschalsteuer, die aber voraussetzen würde, dass der Staat die Lohnsätze beobachten kann.

Die Problematik dieser Modellformulierung sieht man, wenn man die self-selection-constraints weglässt und die erstbeste (allerdings nicht realisierbare) Lösung studiert: sie weist ein mit den Fähigkeiten fallendes Nutzenniveau auf (Mirrlees 1974, Brunner 1989). Es fällt vermutlich schwer, eine solche Umkehr der Nutzenpositionen im Vergleich zur Situation ohne Steuer als eine gute (gerechte) Lösung für die Einkommensteuer zu akzeptieren; erst durch die Hinzunahme der Informationsbeschränkung resultiert eine Situation, bei der die Nutzenpositionen mit den Fähigkeiten steigen, als optimal. Auf einer grundsätzlichen Ebene kann man nun fragen, ob eine Norm und deren Anwendung in einem Modell, die bei einer erstbesten Lösung zu einem inakzeptablen Ergebnis führt, als sinnvoll angesehen werden kann.[39]

B.III.4. Die optimale Einkommensteuer und die Besteuerung von Gütern und Kapitaleinkommen

Nun wenden wir uns dem Verhältnis zwischen direkten und indirekten Steuern zu. Als Unterscheidungsmerkmal zwischen diesen beiden Gruppen dient üblicherweise die Überwälzungsmöglichkeit oder die Möglichkeit zur Berücksichtigung persönlicher Umstände. Dies wird allerdings im vorliegenden Modell nicht explizit modelliert, wir folgen einfach der üblichen Sichtweise und sehen die Steuer auf Arbeitseinkommen als eine direkte Steuer, die Steuern auf Güter, in unserem Modell mit $\hat{\tau}_1, \hat{\tau}_2$ bezeichnet, als indirekte Steuern an.[40]

Es sei, für gegebene $\hat{\tau}_1, \hat{\tau}_2$ der maximale Wert der Zielfunktion bei der Lösung des Zwei-Personen-Optimierungsproblems mit nichtlinearer Einkommensteuer (B.II.5), (B.II.2), (B.II.3), (B.II.4), (B.II.6)

[39] Eine Alternative wäre die Anwendung des Rawls-Kriteriums, das in der Theorie ähnliche Eigenschaften der zweitbesten Lösung wie die utilitaristische Zielfunktion liefert (Atkinson 1973, Itsumi 1974, Brunner 1989), aber zu ausgeprägter Umverteilung (negativer Steuer) führt, andererseits aber auch bei der erstbesten Lösung keine Umkehrung der Nutzenpositionen impliziert. Eine andere Möglichkeit wäre der Rückgriff auf ein Opferprinzip (siehe Berliant and Gouveia 1993), das der Idee eines persönlichen Anspruchs auf das eigene Einkommen näher kommt, wie in der Einleitung erörtert.

[40] Zu dieser Unterscheidung siehe Atkinson 1977.

mit $S(\hat{\tau}_1, \hat{\tau}_2)$ bezeichnet. Mithilfe des Envelope-Theorems kann man die partielle Ableitung von S nach $\hat{\tau}_i$ an der Stelle $\hat{\tau}_1 = \hat{\tau}_2 = 0$ bestimmen. Sie ergibt sich aus den Bedingungen erster Ordnung nach einigen Umformungen (siehe Anhang 3) als

$$\left.\frac{\partial S}{\partial \tau_i}\right|_{\hat{\tau}_1=\hat{\tau}_2=0} = \mu_2 \frac{\partial v^2}{\partial x^1}(c_i^2[1] - c_i^1), \qquad i = 1, 2. \qquad \text{(B.III.3)}$$

Dabei stellt μ_2 den Lagrange-Multiplikator zur self-selection constraint (B.II.3) dar, er ist positiv, wenn wir wie oben annehmen, dass diese Bedingung bindet (vgl. Abbildung 4), also die (erwünschte) Umverteilung von Person zwei zu Person eins dadurch beschränkt ist, dass Person zwei bei einem zu hohen Ausmaß der Umverteilung das der Person eins zugedachte (z^1,x^1)-Bündel vorziehen würde. $\partial v^2 / \partial x^1$ ist ebenfalls positiv, es ist der Grenznutzen des Nettoeinkommens für Person zwei, falls sie das andere Bündel wählt. $c_i^2[1]$ bezeichnet den Konsum des Gutes i der Person zwei in diesem letzteren Fall (bei mimicking).

Die Antwort auf die Frage, ob eine positive Steuer oder eventuell eine Subventionierung $\hat{\tau}_i$ auf ein Gut erwünscht ist oder nicht, gemäß der Zielfunktion, hängt von der Differenz zwischen dem Konsum des Gutes i durch die Person eins, c_i^1, und den Konsum $c_i^2[1]$ dieses Gutes durch Person zwei, wenn sie eins imitiert, ab. Eine intuitive Erklärung dafür liegt in der Überlegung, dass im betrachteten Modell der Effekt der indirekten Steuer auf die Selbstselektionsbedingung (B.II.3) maßgeblich ist. Im Fall $c_i^2[1] > c_i^1$ wird diese Bedingung durch die Einführung von $\hat{\tau}_i$ weniger restriktiv, weil Person zwei, wenn sie Person eins imitiert, durch $\hat{\tau}_i$ mehr verliert als Person eins und daher weniger "schnell" zum Imitieren neigt. M. a. W.: man kann Umverteilung in einem größeren Ausmaß durchführen.

Der Unterschied, der für die Differenz zwischen c_i^1 und $c_i^2[1]$ den Ausschlag gibt, besteht darin, dass Person zwei weniger Arbeitszeit benötigt, um das Bruttoeinkommen z^1 (und somit das Nettoeinkommen x^1) zu erzielen als Person eins. Damit erhalten wir das klassische Resultat (Atkinson and Stiglitz 1976, Deaton 1981): $\hat{\tau}_i = 0$ ist optimal, wenn die (identische) Nutzenfunktion $U(c_1,c_2,L)$ schwach separabel zwischen den Konsumgütern und Freizeit ist. Dann hängt die Aufteilung des Nettoeinkommens x^1 auf die beiden Güter nämlich nicht von der Höhe der Arbeitszeit ab, somit ist in (B.III.3) $c_i^2[1] = c_i^1$ und $\partial S / \partial \hat{\tau}_i = 0$. Liegt schwache Separabilität nicht vor, so steigt der Wert der sozialen Wohlfahrt, wenn die Steuer auf jenes Gut positiv ist, für das

$c_i^2[1] > c_i^1$ gilt, das also ein Komplement zur Freizeit darstellt - wie wir es von der Corlett-Hague-Regel kennen, nun allerdings durch ein Umverteilungsmotiv ergänzt.[41]

Durch dieses Ergebnis erscheint die Begründung für indirekte Steuern zumindest abgeschwächt. Sie erweisen sich bei Vorliegen einer nicht ganz unplausiblen Klasse von Präferenzen als nicht erforderlich, eine einheitliche Steuer im Sinn der existierenden Umsatzsteuer kann schon gar nicht in diesem Modell begründet werden, weil die optimalen Steuersätze ja von der Komplementarität einzelner Güter mit Freizeit abhängen.[42] Das prinzipielle Problem ist allerdings, dass man durch empirische Untersuchungen kaum Aussagen über schwache Separabilität bzw. Substitutionsbeziehungen gewinnen kann (Deaton 1981). Diese Situation ist schon vom Ramsey-Modell bekannt; als wesentliche Schlussfolgerung erscheint, dass die Betrachtung von Personen mit unterschiedlichen Lohnsätzen und damit die Einbeziehung des Verteilungsproblems wenig an den Resultaten zur Güterbesteuerung ändert, wenn es eine optimale nichtlineare Einkommensteuer gibt. Insgesamt ist der Schluss wohl, dass in diesem Modell, wenn sich die Personen in ihren Fähigkeiten zur Erzielung von Arbeitseinkommen unterscheiden, eine Steuer auf dieses zur Umverteilung ausreicht. [43]

Man kann zeigen (Edwards u. a. 1994), dass im Fall von Gütersteuern $\hat{\tau}_1, \hat{\tau}_2$ der effektive marginale Steuersatz (also der Marginalsatz, der sich aus der Kombination von Einkommensteuer und Gütersteuern ergibt) positiv für Person eins und null für Person zwei ist. Der optimale Marginalsatz der Einkommensteuer allein ist dann für Person zwei nicht null (er ist negativ, falls beide Güter besteuert werden).

Eine weitere Interpretation dieses Modells, auf die schon von Atkinson und Stiglitz (1976) hingewiesen wurde, erhält man, wenn man die beiden Güter c_1, c_2 in einem zwei-Perioden-Modell als Gegenwarts- und Zukunftskonsum interpretiert. Dann folgt, dass bei schwacher Separabilität die Einhebung einer Steuer auf Kapitaleinkommen (eine indirekte Steuer auf das Zukunftsgut) nicht optimal ist. Da gerade in diesem Fall eine Aussage über Komplementaritäts- bzw. Substitutionsbeziehungen zu Freizeit praktisch schwer zu finden ist, bleibt die Beurteilung einer Steuer auf Kapitaleinkommen in diesem Modell ungeklärt.[44]

[41] Ein ähnliches Resultat gilt auch im Fall einer optimalen linearen Einkommensteuer, wenn die Präferenzen bezüglich der Güter homothetisch sind (Deaton 1981).
[42] In verschiedenen Arbeiten wurde auch untersucht, unter welcher Bedingung die mit den Steuereinnahmen finanzierten öffentlichen Güter zur Umverteilung eingesetzt werden können, zusätzlich zur optimalen Einkommensteuer. Siehe etwa Christiansen 1981, Boadway and Keen 1993.
[43] Boadway, u. a (1994) zeigen, dass die Besteuerung der Konsumgüter, zusätzlich zu einer optimalen Einkommensteuer, aufgrund von Steuerhinterziehung sinnvoll sein kann. In Naito (1999) kann eine differenzierte Besteuerung der Konsumgüter als Pareto-verbessernd wirken, wenn die beiden Personen unterschiedliche und nicht konstante Grenzproduktivitäten der Arbeit aufweisen.
[44] Zur Diskussion dieser Frage siehe auch Homburg (2000), S. 186.

Offensichtlich liegt ein Mangel jedes Modells in der geringen Zahl von relevanten Aspekten, die es abbildet. In dieser Hinsicht betrifft eine häufig geäußerte Kritik am Mirrlees-Modell die alleinige Einbeziehung von Unterschieden im Lohnsatz, während in der Realität noch eine Vielzahl weiterer Unterscheidungsmerkmale von Bedeutung sind. So liegt es nahe, Unterschiede in Ausstattungsmerkmalen an Gütern bzw. Kapital einzubeziehen, wie dies in jüngeren Arbeiten unternommen wurde (Cremer u.a. 2001, Brunner 1999). Um den Effekt dieser Erweiterung zu sehen, schreiben wir die Budgetbedingung eines Haushalts in der Form

$$(p_1 + \hat{\tau}_1)(c_1^h - \bar{c}_1^h) + (p_2 + \hat{\tau}_2)(c_2^h - \bar{c}_2^h) \leq z - \sigma(z) + (1 - \tau_e)e^h \qquad h = 1,2, \qquad \text{(B.III.2)}$$

wobei $\bar{c}_1^h$, $\bar{c}_2^h$ die Anfangsausstattungen des Haushalts an den beiden Gütern bezeichnen, σ eine beliebige Steuerfunktion in Bezug auf das Bruttoarbeitseinkommen und e^h ein Anfangsvermögen, das als Substitut für Arbeit in der Güterproduktion verwendet werden kann und das mit einem proportionalen Steuersatz τ_e belegt werden kann.[45] Es zeigt sich, dass nicht alle vier Steuern in diesem Modell erforderlich sind; man kann etwa (unter der Annahme $\tau_e < 1$) ein äquivalentes System ohne Steuer auf das Vermögen finden, indem man die beiden Gütersteuern sowie die Einkommensteuer entsprechend erhöht.[46] Im Weiteren sei daher $\tau_e = 0$ gesetzt.

Eine Konsequenz der Existenz von unterschiedlichen Anfangsausstattungen ist, dass die Plausibilität der single-crossing-Bedingung etwas geringer wird. Unter der Annahme eines abnehmenden Grenznutzens des Einkommens kann ein hohes Vermögen von Person zwei dazu führen, dass sie als Kompensation für die Mühe zur Erzielung einer weiteren Einheit Bruttoeinkommen mehr an zusätzlichem Nettoeinkommen benötigt als Person eins. Dann wirkt die Einkommensteuer in die falsche Richtung; dies wird im Folgenden ausgeschlossen.

Wir betrachten zunächst den Fall, dass die Anfangsausstattungen im Form von Vermögen vorliegen, und nehmen an, dass e^1 nicht um sehr viel größer als e^2 ist, sodass im Optimum weiterhin eine Umverteilung von Person zwei zu Person eins erwünscht ist. Dann bleibt die Optimalitätsbedingung (B.III.1) weiterhin gültig, aber sie enthält nun ein weiteres Element: Zusätzlich zur geringeren Arbeitszeit, die Person zwei aufwendet, wenn sie Person eins imitiert, hat sie nun auch ein anderes Vermögen. Bei schwacher Separabilität zwischen Konsum und Freizeit hat die unterschiedliche Arbeitszeit keinen Effekt. Für den Einfluss des Vermögens aber gilt: wenn es sich um ein normales Gut handelt, so wird es von der Person mit höherem Vermögen

---

[45] Da e als Anfangsausstattung fix gegeben ist, stellt τ_e eine Pauschalsteuer dar, aber nicht differenziert nach Personen.

[46] Klarerweise lassen sich umgekehrt zwei identische Gütersteuern $\hat{\tau}_1 = \hat{\tau}_2$ durch eine Erhöhung der Einkommen- und der Vermögensteuer ersetzen.

in größerem Ausmaß konsumiert und daraus folgt: in Abhängigkeit von $e^2 \gtrless e^1$ soll das Gut i besteuert bzw. subventioniert werden.

Mit diesem Resultat wird der Ausstattungs-Umverteilungsaspekt der indirekten Besteuerung beschrieben. Wenn Ausstattungen und Lohnsätze positiv korreliert sind, ergänzt jede indirekte Steuer (und auch eine Steuer auf Kapitaleinkommen) die optimale nichtlineare Einkommensteuer, und zwar umso mehr, je stärker der Zusammenhang zwischen Anfangsausstattung und Konsum des betreffenden Gutes ist. Ein ähnliches Resultat ergibt sich, wenn man c_1^h und c_2^h als Güter in verschiedenen Perioden interpretiert und e^h als Anfangskapital, das zusammen mit der Ersparnis zur Produktion in der zweiten Periode eingesetzt werden kann. In einem Modell mit nicht korrelierten Fähigkeiten und Anfangsausstattungen erhöht die Einführung einer Steuer auf Kapitaleinkommen in der zweiten Periode, zusätzlich zu einer linearen Steuer auf Arbeitseinkommen, die soziale Wohlfahrt (Vogelsang 2000).

Der gleiche Effekt zeigt sich auch, wenn die Anfangsausstattung aus Konsumgütern besteht. Allerdings tritt hier die Komplikation auf, dass eine Steuer auf ein Gut auch deren Wert in der Anfangsausstattung erhöht, was in die Gegenrichtung wie der erwünschte Umverteilungseffekt wirken kann (Brunner 1999).

Saez (2002b) hat die Erwünschtheit der Güterbesteuerung, zusätzlich zu einer optimalen nichtlinearen Einkommensteuer, in einem sehr allgemeinen Modell untersucht, das mehrere Dimensionen zulässt, in denen sich die Personen unterscheiden. Insbesondere betrachtet er Unterschiede in den Präferenzen (die ja im Mirrlees-Modell als identisch angenommen werden) und zeigt (die Überlegung beruht auf einer Formel analog zu (B.III.3)), dass die Besteuerung eines Gutes, für das Personen mit höherem Einkommen eine stärkere Präferenz aufweisen, aus Verteilungsgründen sinnvoll sein kann. Dies gilt auch für die Besteuerung von Kapitaleinkommen in einem Zwei-Perioden-Modell, wenn die Sparneigung mit dem Lohnsatz steigt.

Insgesamt wirft die Hinzunahme unterschiedlicher Anfangsausstattungen neues Licht auf die Diskussion über die Rolle indirekter Steuern. Im Gegensatz zur früher oft betonten regressiven Wirkung können sie auch eine Ergänzung zur durch eine Einkommensteuer möglichen Umverteilung bringen, oder es wird dies durch eine Steuer auf das Anfangsvermögen direkt bewirkt.

Unbefriedigend bleibt bei diesem Modell, dass die Anfangsvermögen (in Form von Gütern oder Kapital) einfach als existierend angenommen werden, sodass ihre Besteuerung Pauschalcharakter hat. Offensichtlich kann man sie am ehesten als Erbschaften interpretieren, wie das Cremer u. a. (2001) tun. Für eine umfassendere Betrachtung erscheint daher ein dynamischer und intergenerativer Kontext geeigneter, dem wir uns im nächsten Abschnitt zuwenden, wobei zunächst die bekanntesten Modelle zur Besteuerung von Kapitaleinkommen diskutiert werden.

C. Dynamische Modelle der optimalen Besteuerung

C.I. Unendlicher Zeithorizont der Personen

Wir gehen nun der Frage nach der optimalen Ausgestaltung der Steuern auf Arbeits- und Kapitaleinkommen in Modellen nach, die intertemporale Entscheidungen der Haushalte berücksichtigen. Offensichtlich ist die Übertragung des mikroökonomischen Rationalkalküls auf den intertemporalen Kontext nicht unproblematisch, wegen der mit langen Zeiträumen unweigerlich verbundenen Unwägbarkeiten. Trotzdem folgen wir dem üblichen Vorgehen, in diesen Modellen die Konsequenzen der Steuern zu erörtern.

Als erstes wenden wir uns jener Modellvariante zu, die den Haushalten eine unendlich lange Voraussicht und Planungskapazität (ein Denken in "Dynastien") unterstellt.[47] Gegeben seien wieder zwei Haushalte h = 1, 2, deren intertemporale Präferenzen über Konsum und Arbeitszeit in den Perioden t = 0, 1, 2, ... (bis unendlich) sich durch eine additive Nutzenfunktion darstellen lassen:

$$\sum_{t=0}^{\infty} \beta^t U^h(c_t^h, L_t^h). \tag{C.I.1}$$

Dabei ist U^h eine übliche strikt konkave Nutzenfunktion bezüglich Konsum und Arbeit in jeder Periode. U^h kann für die beiden Haushalte unterschiedlich sein, aber ihre subjektiven Diskontfaktoren $\beta < 1$ (bzw. die Zeitpräferenzraten) stimmen überein. Die Budgetbedingung eines Haushalts in Periode t lautet

$$c_t^h + k_{t+1}^h + b_{t+1}^h = [(1-\tau_{wt}) w_t^h L_t^h + R_{kt} k_t^h + R_{bt} b_t^h], \qquad t = 0,1,2,..., \tag{C.I.2}$$

wobei $R_{kt} \equiv 1 + (1-\tau_{rt})(r_t - \delta)$ den Rückfluss aus einer Kapitaleinheit (Zinssatz r_t), nach Abzug der Steuer τ_{rt} und der Abschreibung δ, in Periode t beschreibt. k_t^h bezeichnet den Kapitalbestand in Periode t im Besitz der Person h, wobei der Anfangsbestand k_0^h, h = 1, 2 vorgegeben ist. b_t^h bezeichnet den Bestand an Krediten (für jeweils eine Periode) an den Staat, den Person h in Periode t hält, R_{bt} die Rückzahlung plus Verzinsung pro Einheit (b_0^h ist gegeben). τ_{wt} bezeichnet eine proportionale Steuer auf Arbeitseinkommen in Periode t. Für die Haushaltsentscheidung wird

[47] Die Darstellung folgt weitgehend den Arbeiten von Chari und Kehoe (1999) und Atkeson, Chari und Kehoe (1999).

perfekte Voraussicht bezüglich aller Größen, vor allem auch der Steuersätze, unterstellt.[48] Dann ergeben sich die Bedingungen erster Ordnung für die individuelle Optimierung:

$$\partial U^h / \partial c_t^h = R_{kt+1} \beta \partial U^h / \partial c_{t+1}^h \tag{C.I.3}$$

$$\partial U^h / \partial L_t^h = -(1-\tau_{wt}) w_t^h \partial U^h / \partial c_t^h . \tag{C.I.4}$$

Die Produktion in jeder Periode kann durch die linear-homogene Funktion $F(k_t, L_t^1, L_t^2)$ beschrieben werden, wobei keine vollkommene Substituierbarkeit der von den beiden Personen angebotenen Arbeit vorliegen muss.[49] Wettbewerb auf den Märkten begründet die Entlohnung der Faktoren mit ihrem Grenzprodukt ($r_t = \partial F / \partial k_t, w_t^h = \partial F / \partial L_t^h$).

Die Aufgabe des Staates besteht in der Wahl geeigneter Steuern τ_{wt}, τ_{rt} und öffentlicher Verschuldung, wobei die Zielsetzung durch eine intertemporale soziale Wohlfahrtsfunktion

$$\rho^1 \sum_{t=0}^{\infty} \beta^t U^1(c_t^1, L_t^1) + \rho^2 \sum_{t=0}^{\infty} \beta^t U^2(c_t^2, L_t^2), \tag{C.I.5}$$

mit Gewichten $\rho^1 + \rho^2 = 1$ beschrieben wird. Die Budgetrestriktion des Staates lautet

$$b_{t+1}^1 + b_{t+1}^2 = R_{bt}(b_t^1 + b_t^2) + g_t - \tau_{wt}(w^1 L_t^1 + w^2 L_t^2) - \tau_{rt}(r_t - \delta)(k_t^1 + k_t^2), \tag{C.I.6}$$

wobei g_t den exogen gegebenen öffentlichen Ressourcenanspruch in Periode t bezeichnet. Wie bei allen Problemen der optimalen Besteuerung ist zu berücksichtigen, dass der Staat die Festlegung von Konsum und Arbeitszeit, die in die soziale Wohlfahrtsfunktion einfließen, nicht selbst vornimmt; sie ergeben sich durch die Entscheidungen der Personen selbst, gegeben die dafür notwendigen Parameter. Während wir dies in Abschnitt B.I. durch die Verwendung der *dualen* (indirekten) Nutzenfunktion berücksichtigt haben, gehen wir nun gemäß dem *primalen* Ansatz (ähnlich wie beim Mirrlees-Modell) vor. Wir fragen: welche Bedingungen müssen Werte c_t^h, L_t^h erfüllen, damit es Steuersätze gibt, bei denen sie als Wettbewerbsgleichgewichte aus den individuellen Entscheidungen, resultieren. Dabei wird unterstellt, dass der Staat eine sogenannte *commitment-Technologie* besitzt, mit deren Hilfe er sich selbst an die in Periode null für alle

[48] Chari und Kehoe 1999 analysieren eine Erweiterung des Modells um stochastische Schocks.
[49] Zusammen mit der Annahme nicht-identischer Präferenzen U^h ist daher mehr Unterschiedlichkeit als nur in der Arbeitseffizienz, wie im Mirrlees-Modell, zugelassen. Allerdings ist die Besteuerung der Arbeitseinkommen auf eine proportionale Steuer beschränkt.

Zukunft festgelegten Steuersätze binden kann, so dass sie tatsächlich eine glaubwürdige Grundlage für individuelle Entscheidungen darstellen.

Man kann zeigen (Chari und Kehoe 1999), dass die folgenden *Implementierbarkeitsbedingungen*

$$\sum_{t=0}^{\infty}\beta^t\left(c_t^h\partial U^h/\partial c_t^h + L_t^h\partial U^h/\partial L_t^h\right) = \partial U^h/\partial c_0^h\left(R_{k0}k_0^h + R_{b0}b_0^h\right) \qquad h = 1, 2, \qquad \text{(C.I.7)}$$

$$MRS^1_{c_t^1 c_{t+1}^1} = MRS^2_{c_t^2 c_{t+1}^2} \qquad t = 0, 1, 2, ... \qquad \text{(C.I.8)}$$

$$MRS^1_{L_t^1 c_1^t} / MRS^2_{L_t^2 c_t^2} = MRtS_{L_t^1 L_t^2},\ t = 0,1,2, ... \qquad \text{(C.I.9)}$$

zusammen mit der gesamtwirtschaftlichen Ressourcenbeschränkung

$$c_t^1 + c_t^2 + k_{t+1}^1 + k_{t+1}^2 + g_t = F(k_t^1 + k_t^2, L_t^1, L_t^2) + (1-\delta)k_t \qquad \text{(C.I.10)}$$

hinreichend und notwendig dafür sind, dass sich c_t^h, L_t^h aus individuellen Entscheidungen bei geeigneten Steuersätzen ergeben. Dabei garantiert (C.I.8), dass sich beide Personen dem gleichen Steuersatz auf Kapitaleinkommen gegenüber sehen, (C.I.9) garantiert gleichen Steuersatz auf Arbeitseinkommen ($MRtS_{L_t^1 L_t^2}$ ist die Grenzrate der technischen Substitution von L_t^1 und L_t^2 gemäß der Produktionsfunktion). (C.I.7) sind die generellen Implementierbarkeitsbedingungen für die intertemporale Optimierung des Staates. Maximierung von (C.I.5) unter den Nebenbedingungen (C.I.7) - (C.I.10) ergibt die gesuchte optimale Allokation, die zusammen mit den zugehörigen Steuersätzen auf Arbeits- und Kapitaleinkommen sowie der Verschuldung ein sogenanntes *Ramsey*-Gleichgewicht darstellen. Die Bedingungen erster Ordnung, eingeschränkt auf den steady-state (falls ein solcher existiert), lauten für die Kapitalallokation

$$-\beta\Gamma + 1 = \beta(1-\delta+\partial F/\partial k), \qquad \text{(C.I.11)}$$

dabei ist Γ ist eine Konstante. Vergleicht man dies mit der Bedingung (C.I.3) für die individuell optimale intertemporale Allokation im steady-state (d. h. $c_t^h = c_{t+1}^h$), nämlich (eingesetzt für R)

$$1 = \beta(1 + (1-\tau_r)(r-\delta)), \qquad \text{(C.I.12)}$$

so folgt (man beachte $\partial F/\partial k = r$), dass im Fall $\Gamma = 0$ ein optimaler steady-state nur dann mit der individuellen Optimierung vereinbar ist, also ein Ramsey-Gleichgewicht darstellt, wenn die Steuer τ_r auf Kapitaleinkommen gleich null ist. Dabei gilt $\Gamma = 0$ genau dann, wenn die Grenzrate der Substitution zwischen L^1 und L^2 unabhängig vom eingesetzten Kapital ist, also die Produktionsfunktion schwach separabel zwischen den beiden Typen von Arbeit und Kapital.[50]

Dieses Resultat geht auf Chamley (1986) und Judd (1985) zurück. Es gilt für eine beliebige Verteilung der Gewichte ρ^1, ρ^2, sogar für den Fall, dass eine Person nur Arbeit anbietet, die andere nur Kapital und letztere das Gewicht null in der Zielfunktion bekommt. Für Pfade außerhalb des steady-states kann man zeigen, dass die Besteuerung des Kapitaleinkommens in den Anfangsperioden beträchtlich sein kann (der Anfangskapitalbestand ist ja gegeben, die Steuer auf das darauf entfallende Einkommen ist also eine Pauschalsteuer), aber dann auf null fällt.[51] Eine Begründung für das Verschwinden der Kapitaleinkommensteuer im optimalen steady-state liegt darin, dass ein konstanter, von null verschiedener Steuersatz eine gegen unendlich gehende Verzerrung zur Folge hätte. Dies sieht man, wenn man für konstantes k und r den effektiven Gegenwartspreis für c_t^h in anderer Form schreibt (vgl. Bernheim 2002), nämlich

$$\frac{1}{(1+(1-\tau_r)r)^t} = \frac{1}{(1+r)^t}(1+\varphi_t)\,, \tag{C.I.13}$$

wobei φ_t den auf den Gegenwartspreis $1/(1+r)^t$ des Konsums in Periode t aufgeschlagenen Steuersatz bezeichnet, der durch eine konstante Steuer τ_r induziert wird, die von Kapitaleinkommen in jeder Periode eingehoben wird. Es gilt die Rekursionsformel (setze $\varphi_0 = 0$)

$$1+\varphi_t = (1+\varphi_{t-1})\frac{1+r}{1+(1-\tau_r)r}, \tag{C.I.14}$$

was bedeutet, dass der Steuersatz φ_t gegen unendlich geht, wenn $\tau_r > 0$, weil dann der Quotient $(1+r)/(1+(1-\tau_r)r)$ größer als eins ist. Es erscheint plausibel, dass ein gegen unendlich gehender Steuersatz nicht optimal sein kann.[52]

[50] Dies ist insbesondere dann der Fall, wenn L^1 und L^2 perfekte Substitute darstellen, wie im Mirrlees-Modell.
[51] Damit ergibt sich offensichtlich ein Problem der Zeitinkonsistenz: In jeder Periode kann der Staat aufs Neue den gegebenen Kapitalstock zu besteuern versuchen, wenn die commitment-Technologie nicht funktioniert. Eine zeitlich konsistente Lösung wird u. a. in Benhabib und Rustichini (1997) und Krusell (2002) erörtert.
[52] Von Correia (1996) wurde gezeigt, dass die Nichtbesteuerung von Kapitaleinkommen nicht länger optimal ist, wenn es einen weiteren Produktionsfaktor gibt, der unbesteuert bleibt. Damit ergibt sich eine Parallelität zu den Aussagen über die Optimalität von Produktionseffizienz (siehe Fußnote 12), für die Ähnliches gilt (Munk 1980).

Für eine um Humankapital erweiterte Version des Modells zeigten Jones u. a. (1997), dass im optimalen steady-state auch die Steuer auf Arbeitseinkommen null ist, so dass nur eine Konsumbesteuerung übrig bleibt. Andererseits wurden in der jüngeren Zeit in verschiedenen Beiträgen andere Erweiterungen des Modells vorgestellt, die eine Besteuerung der Kapitaleinkommen im optimalen steady-state als Lösung ergeben, etwa wegen Kreditbeschränkungen (Chamley 2001, Aiyagari 1995) oder durch die Forderung nach jährlichem Budgetausgleich (Lansing 1999).

Generell gilt, dass die Annahmen, auf denen diese Art von Modellen beruht, doch erheblich sind, vor allem die unendliche Voraussicht der Individuen und ihre darauf aufbauende Rationalität der intertemporalen Entscheidung, sowie die Existenz einer commitment-Technologie.

C.II. Überlappende Generationen

Als Alternative zur offensichtlich unrealistischen Annahme einer rationalen Planung bei vollständiger Voraussicht über einen unendlich langen Zeithorizont bietet sich das Modell überlappender Generationen an, das von Diamond (1965) in die Wachstumstheorie eingeführt wurde. Im einfachsten Fall nehmen wir an, eine Person lebt für zwei Perioden und in jeder Periode tritt eine neue Generation, bestehend aus zwei Personen mit unterschiedlichen Arbeitsproduktivitäten $w^1 < w^2$, in die Ökonomie ein. Kapitalbildung erfolgt durch Sparen im Lebenszyklus: die Personen arbeiten und sparen (bilden einen Kapitalstock) in der ersten Periode, sie befinden sich im Ruhestand und entsparen (verkaufen den Kapitalstock) in der zweiten Periode.[53]

Wir betrachten zum Studium der Eigenschaften optimaler Steuern eine unmittelbare Erweiterung des in B.II. erörterten Mirrlees-Modells. Die beiden Güter werden jetzt mit c_{1t}^h, c_{2t+1}^h bezeichnet und sind als Konsum eines in Periode t eintretenden Haushalts h in seiner ersten bzw. zweiten Lebensphase zu interpretieren. Zusätzlich zu den Paaren (z^h, x^h), h = 1, 2, legt der Staat den Nettozinssatz r_t in Periode t fest sowie einen pauschalen intergenerativen Transfer b_t, von einer aktiven Person zu einer in der Altersperiode (vom Bevölkerungswachstum sehen wir aus Gründen der Einfachheit ab.) Somit ergibt sich die direkt-indirekte Nutzenfunktion für Person h als $v^h(z_t^h, x_t^h, r_{t+1}, b_t, b_{t+1}) \equiv \max\{(U(c_t^h, c_{2t+1}^h, z^h / w^h) / c_{1t}^h + c_{2t+1}^h/(1+r_{t+1}) \leq x_t^h - b_t + b_{t+1}/(1+r_{t+1})\}$ und damit die intertemporale Zielfunktion des Planers

[53] Vorläufig sehen wir von Vererbung ab. Im Modell von Barro (1974) sind die Generationen über Vererbung mit den Nachfolgern verbunden, wodurch sich dann indirekt wieder ein unendlicher Planungshorizont ergibt. Offensichtlich setzt dies eben die Kenntnis der Nutzenpositionen zukünftiger Generationen (also vollkommene Voraussicht) voraus.

$$\max_{z_t^1, z_t^2, x_t^1, x_t^2, r_t, b_t} \sum_{t=0}^{\infty} \gamma^t (\rho^1 v^1(\cdot) + \rho^2 v^2(\cdot)), \qquad \text{(C.II.1)}$$

wobei $\gamma < 1$ den Diskontfaktor bezeichnet, mit dem das Gewicht jeder folgenden Generation geringer wird. Der Kapitalstock (die Ersparnis) k_t^h, den Person h in Periode t - 1 für die Produktion in Periode t bildet, lässt sich als

$$k_t^h(z_{t-1}^h, x_{t-1}^h, r_t, b_{t-1}, b_t) = (c_{2t}^h(z_{t-1}^h, x_{t-1}^h, r_t, b_{t-1}, b_t) - b_t)/(1 + r_t) \qquad \text{(C.II.2)}$$

schreiben. Damit ergibt sich die aggregierte Ressourcenrestriktion in Periode t in der Form

$$x_t^1 + x_t^2 + [k_t^1(\cdot) + k_t^2(\cdot)](1 + r_t) + g_t \le F(k_t^1(\cdot) + k_t^2(\cdot), z_t^1 + z_t^2) + (1 - \delta)(k_t^1(\cdot) + k_t^2(\cdot)) \text{ , t = 0,1,2, ...,} \qquad \text{(C.II.3)}$$

wobei $F(k_t,z_t)$ die linear-homogene Produktionsfunktion (mit aggregiertem Kapital $k_t = k_t^1 + k_t^2$ und effektiver Arbeit $z_t = z_t^1 + z_t^2$ als Argumente) bezeichnet, δ die Abschreibungsrate. g_t, t = 0,1,2, ... ist wieder der exogen gegebene Ressourcenanspruch des Staates, $k_t^h(1 + r_t)$ stellt das Entsparen der Rentner dar, ihr Konsum ergibt sich als $c_{2t}^h = k_t^h(1 + r_t) + b_t$, wobei b_t aber gleichzeitig den Konsum der Erwerbstätigen reduziert und daher in (C.II.3) nicht aufscheint.

Zu beachten ist, dass zwar x_t^h wieder das Nettoeinkommen beschreibt, aber z_t^h nun nicht mehr - wie im statischen Modell - das Bruttoeinkommen. Letzeres ergibt sich durch die Bewertung mit dem Grenzprodukt der Arbeit als $z_t^h \partial F / \partial z_t$.[54] Man überlegt sich leicht, dass wiederum die bezüglich z_t^h, x_t^h definierten Selbstselektionsbeschränkungen die Existenz einer Steuerfunktion garantieren, bei der beide Personen die ihnen zugedachten Bündel selbst wählen.[55] Zur Vereinfachung gehen wir entsprechend den in (B.II.) angestellten Schlussfolgerungen davon aus, dass nur jene für Person zwei bindet, also berücksichtigt werden muss:

$$v^2(z_t^2, x_t^2, r_{t+1}, b_t, b_{t+1}) \ge v^2(z_t^1, x_t^1, r_{t+1}, b_t, b_{t+1}), \qquad t = 0,1,2, \ldots \qquad \text{(C.II.4)}$$

[54] Anders formuliert: Das Grenzprodukt der effektiven Arbeit wurde in B.II. als konstant mit dem Wert 1 angenommen.
[55] Allerdings ist dabei eine Modifikation der im statischen Modell an Hand der Abbildung 3 vorgenommenen Überlegung erforderlich, weil die Steuerfunktion ja auf dem Bruttoeinkommen $z_t^h \partial F / \partial z_t$ und nicht auf der effektiven Arbeitsleistung z_t^h definiert ist, die im vorliegenden Modell die staatliche Entscheidungsvariable darstellt.

Maximierung von (C.II.1) unter Beachtung von (C.II.3) und (C.II.4) sowie der Nichtnegativitätsbedingungen liefert Bedingungen erster Ordnung für z_t^h, x_t^h, somit für c_{1t}^h, c_{2t+1}^h, h = 1,2, t = 0,1,2, ... [56] und ebenso für die Nettozinssätze r_t und die intergenerativen Transfers b_t, t = 0,1,2, ...

Die Differenz zwischen dem vom Staat (optimal) festgelegten Nettozinssatz r_t, der für die Haushalte relevant ist, und dem sich auf einem kompetitiven Kapitalmarkt ergebenden Bruttozinssatz $\partial F / \partial k_t - \delta$, dem Grenzertrag des Kapitals unter Berücksichtigung der Abschreibung, lässt sich als die vom Staat auferlegte Steuer auf Kapitaleinkommen interpretieren. Wenn man steady-states betrachtet, so kann man zeigen (zur Herleitung siehe Brunner 2003), dass sich für den Fall keiner Kapitalertragsteuer, d. h. $r = \partial F / \partial k - \delta$, die modifizierte *goldene Regel* als optimale Lösung für die Kapitalbildung ergibt:

$$1/\lambda = \partial F / \partial k + 1 - \delta = 1 + r \ .$$

Hiefür ist das Instrument der intergenerativen Transfer wesentlich, mit dessen Hilfe der Staat die dezentrale Entscheidung über die Kapitalallokation in der gewünschten Weise beeinflussen kann (Pestieau 1974)[57].

Ob nun die optimale steady-state-Lösung eine Kapitalertragsteuer erfordert oder nicht, hängt schließlich vom Vorzeichen von

$$\left.\frac{\partial L}{\partial r_t}\right|_{r_t = \partial F / \partial k_t - \delta} = \mu_{t-1} \frac{\partial v^2}{\partial x_t^1} \frac{1}{1 + r_t} (k_t^1 - k_t^2[1]) \qquad \text{(C.II.5)}$$

ab, wobei L die Lagrangefunktion zum obigen Maximierungsproblem bezeichnet. Diese Formel ist ganz analog zu interpretieren wie (B.III.1) im statischen Modell, wobei nun $k_1^2[1]$ die Kapitalbildung (Ersparnis) der Person zwei ist, wenn sie das der Person eins zugedachte Bündel wählt. Somit gilt wieder: Bei schwacher Separabilität der Präferenzen zwischen Konsum in beiden Perioden und Freizeit ist der Klammerausdruck in (C.II.5) null und der optimale steady-stae weist genau die

[56] c_{20}^h, h = 1,2 wird als gegeben unterstellt.

[57] Der intergenerative Transfer kann als Umlage im Rahmen der Alterssicherung interpretiert werden. Er kann auch als öffentliche Kreditaufnahme bzw. -rückzahlung modelliert werden (wobei die Zinszahlung des Staates durch die Steuer von den Aktiven eingehoben wird), das Ergebnis gilt unverändert. Anzumerken ist, dass im Fall der dynamischen Effizienz (Knappheit des Kapitals) der Transfer von den Pensionisten zu den Erwerbstätigen gehen muss, um zum optimalen steady-state zu gelangen (siehe u. a. Brunner 1990).

gleichen Eigenschaften auf, die wir vom statischen Modell her kennen, insbesondere ist eine Steuer auf Kapitaleinkommen nicht optimal. Liegt schwache Separabilität nicht vor, so gilt wieder die Corlett-Hague Überlegung, dass eine Steuer auf Kapitaleinkommen (d. h. $r < \partial F / \partial k - \delta$) den Wert der sozialen Wohlfahrtsfunktion erhöht, wenn $k_t^2[1] > k_t^1$ (dann ist die rechte Seite von (C.II.5) negativ; beachte, dass alle anderen Größen in (C.II.5) positiv sind), wenn also der Zukunftskonsum ein Komplement zur Freizeit ist, andernfalls ist eine Subvention sinnvoll.[58] Ordover und Phelps (1979) haben dies für eine kontinuierliche Verteilung der Fähigkeiten gezeigt.[59]

Wenn das Instrument des intergenerativen Transfers nicht (ausreichend) zur Verfügung steht, sind gemäß der Wachstumsperspektive die Steuern auf Arbeits- und Kapitaleinkommen im Hinblick auf ihren Effekt auf die Kapitalbildung zu wählen. Atkinson und Sandmo (1980) betrachten ein Modell mit überlappenden Generationen, identischen Individuen und linearen Steuern und zeigen, dass außerhalb des steady-states die nichtkompensierten Reaktionen relevant sind, im Gegensatz zu den kompensierten Reaktionen (Substitutionseffekten), die im statischen Ramsey-Modell bzw. im steady-state ausschlaggebend sind. Sie geben ein Beispiel, dass in diesem Fall eine positive Steuer auf Kapitaleinkommen optimal sein kann, wenn der Kapitalstock auf einem Anpassungspfad geringer als im optimalen steady-state ist.

Zu betonen ist, dass sich die Resultate für das statische Modell bzw. für überlappende Generationen im steady-state wesentlich von jenen für die Modelle mit unendlichem Zeithorizont unterscheiden. Bei letzteren ist es die gegen unendlich gehende Verzerrung der zukünftigen Preise, die eine Besteuerung von Kapitaleinkommen im steady-state gänzlich unattraktiv macht. Im statischen Modell hängt der positive oder negative Effekt einer Besteuerung der Kapitaleinkommen dagegen von der Art der Präferenzen ab. Dafür wäre im Prinzip eine Abklärung durch empirische Untersuchungen erforderlich. Allerdings handelt es sich - wie schon früher erörtert - um kaum beobachtbare Größen (kompensierte Kreuzpreiseffekte), die zudem schon durch die Vorgabe funktionaler Formen determiniert werden.

Wenn man die Realitätsnähe erörtert, wird man vielleicht geneigt sein, die durchaus schon anspruchsvolle Annahme der rationalen Spar- und Konsumentscheidung über den Lebenszyklus eher zu akzeptieren als jene der perfekten Voraussicht über eine unendlichen Zeithorizont. Immerhin weist die letztere Modellformulierung, wenn man sie als für das Verhalten von Dynastien passend interpretiert, die Besonderheit auf, dass sie die Weitergabe des Kapitalstocks an die

[58] Person zwei kann die effektive Arbeit z^1 bei geringerer Arbeitszeit z^1/w^2 anbieten als Person eins (Arbeitszeit z^1/w^1).
[59] Erosa und Gervais (2002) betrachten ein Lebenszyklusmodell, in dem identische Personen in mehreren Perioden Arbeit anbieten. Sie zeigen, dass eine Steuer auf Kapitaleinkommen auch im steady-state optimal ist, wenn die Steuer auf Arbeitseinkommen nicht altersbezogen ist.

Folgegenerationen einbezieht, was im bisher betrachteten Modell mit überlappenden Generationen nicht der Fall ist.

Wenn man Erbschaften in das Modell überlappender Generationen explizit einbezieht, [60] so liegt es nahe, sie auch als einen möglichen Ansatzpunkt für die Besteuerung, als Teil eines optimalen Steuersystems zu sehen. Dies lässt sich mit der Idee der Chancengleichheit begründen, die eine Angleichung der Anfangsausstattung an Kapital innerhalb jeder Generation nahe legt. Man kann zeigen (Brunner 1997), dass die Einhebung einer Erbschaftsteuer, zusätzlich zu einer optimalen nichtlinearen Einkommensteuer, aus Verteilungsgründen sinnvoll ist, wenn Erbschaften und Fähigkeiten positiv korreliert sind und in der sozialen Wohlfahrtsfunktion hinreichendes Gewicht auf Umverteilung nach unten gelegt wird.

In einem ähnlichen Modell haben Boadway u. a. (2000) die Funktion einer Steuer auf Kapitaleinkommen als Ersatz für eine Erbschaftsteuer begründet, falls letztere nicht im erwünschten Ausmaß eingesetzt werden kann (weil es Wege zu ihrer - teilweisen - Umgehung gibt). Sie betrachten zwei Gruppen von Personen mit unterschiedlichen Lohnsätzen. Jede Person plant den Konsum über zwei Perioden und hinterlässt mit einer bestimmten Wahrscheinlichkeit eine Erbschaft (oder nicht), die in das Budget der Nachkommen bzw. in deren Kapitalstock einfließt. Die Verteilung der Erbschaften und der Fähigkeiten ist voneinander unabhängig. Wie im obigen Modell kann die Regierung öffentliche Verschuldung einsetzen, um die Ökonomie auf den optimalen Entwicklungspfad gemäß der modifizierten goldenen Regel zu bringen. Boadway u. a. zeigen in diesem Modell, dass es den Zielfunktionswert (es wird eine intertemporale utilitaristische soziale Wohlfahrtsfunktion angenommen) erhöht, wenn zusätzlich zu einer optimalen nichtlinearen Einkommensteuer noch eine lineare Steuer auf Kapitaleinkommen eingehoben wird - auch wenn die (identischen) Präferenzen der Personen schwach separabel zwischen dem Konsum in beiden Perioden und Freizeit sind. Die Kapitaleinkommensteuer wirkt als Ersatz für eine Erbschaftsteuer (die im vorliegenden Modell eine Pauschalsteuer wäre).

In einem Lebenszyklus-Modell lässt sich die Budgetbeschränkung einer Person mit einer Lebensdauer von t_m Perioden auf folgende Weise darstellen:

$$I + \sum_{t=1}^{t_m} \frac{wL_t}{(1+r)^{t-1}} = \sum_{t=1}^{t_m} \frac{c_t}{(1+r)^{t-1}} + \frac{B}{(1+r)^{t_m}} \qquad \text{(C.II.6)}$$

[60] Zur Diskussion, welcher Anteil an der Kapitalbildung durch das Sparen im Lebenszyklus bzw. durch Vererbung erklärt werden kann, siehe u. a. Bernheim 2002.

Dabei bezeichnet I eine erhaltene (zu Beginn des Lebens) und B eine am Lebensende hinterlassene Erbschaft. Offensichtlich kann eine Besteuerung sowohl auf der Einkommensseite als auch auf der Ausgabenseite ansetzen, die Beobachtbarkeit der jeweiligen Größen bedingt das Ausmaß, in dem eine Besteuerung möglich ist. Kapitaleinkommen treten in (C.II.6) nicht explizit auf. Im eben besprochenen Modell von Boadway u. a (2000) wird deren Besteuerung durch die zu geringe Erfassbarkeit der Erbschaften begründet. In der Realität gibt es vermutlich eine Reihe weiterer Größen, die aus Gerechtigkeits- oder Effizienzgründen als Anknüpfungspunkte für Steuern dienen könnten, aber nicht direkt beobachtet werden können. Atkinson und Sandmo 1980 erwähnen etwa die Vermutung, dass die Höhe der erzielbaren Kapitalerträge von der Höhe des Vermögens abhängt, z. B. wegen Fixkosten der Veranlagung. Eine andere Möglichkeit wären Unterschiede in der Lebenserwartung, die für die Höhe der Auszahlungen aus Pensionsverträgen bedeutsam sind. Eine "ersatzweise" Besteuerung der Kapitaleinkommen kann auch aus diesen oder ähnlichen Gründen sinnvoll sein.

D. Abschließende Bemerkungen

Die Modelle, die in diesem Beitrag besprochen wurden, geben weitgehend keine detaillierten Antworten auf die Frage, wie die Besteuerung aussehen soll, etwa keinen genauen Verlauf des Einkommensteuertarifs oder keine genaue Höhe der Steuersätze auf Konsumgüter.[61] Für deren nähere Bestimmung müsste man konkrete Informationen über Präferenzen (z.B. kompensierte oder unkompensierte Angebots- und Nachfrageelastizitäten) bzw. Nutzenfunktionen (Grenznutzen des Einkommens) besitzen bzw. müsste die normative Vorgabe (die soziale Wohlfahrtsfunktion) fixiert sein. Es erscheint - wie mehrfach betont - ziemlich ausgeschlossen, dass empirische Untersuchungen jemals eine hinreichende Klarheit über die benötigten Größen liefern werden[62], daher wird die Steuerpolitik stets durch die normativen Positionen der Entscheidungsträger geprägt sein, die ihre Entscheidung mit mehr oder weniger Kenntnis der relevanten ökonomischen Modelle treffen.

In manchen Fällen geben Modelle jedoch sehr dezidierte Hinweise, etwa bezüglich des Spitzen-Grenzsteuersatzes für Arbeitseinkommen oder bezüglich der Kapitaleinkommensbesteuerung. Eine entscheidende Frage ist dann: soll man sie akzeptieren? Mit anderen Worten: stellen diese Modelle ein hinreichend verlässliches Abbild der ökonomischen Realität dar, dass es gerechtfertigt ist, weitreichende wirtschaftspolitische Schlussfolgerungen zu ziehen? Wie wir gesehen haben, gibt es immer auch Varianten dazu, die zu ganz anderen Aussagen führen– tendenziell gilt, dass das Abgehen von der Annahme perfekter Rationalität und effizient funktionierender Märkte die dezidierten Schlussfolgerungen untergräbt. Tatsächlich dürfte aber die Realität durch viele Imperfektionen wie unvollständige Märkte, Kreditrestriktionen, Informationsasymmetrien, mangelnden Wettbewerb, ... gekennzeichnet sein, und insbesondere auch durch Irrationalität der Haushaltsentscheidungen. Gerade letztere ist jedoch nur schwer in die üblichen wohlfahrtsökonomischen Modelle zu integrieren, welche von individueller Rationalität und Konsumentensouveränität ausgehen, um der Problematik paternalistischen Verhaltens des Staates zu entgehen.

Das Problem dabei ist, dass es kein anwendbares Kriterium dafür gibt, ob ein Modell als zutreffend angesehen werden kann oder nicht – empirische Tests erfolgen eher selten, und meistens werden die Modelle (wie etwa das Standardmodell der rationalen Haushaltsentscheidung) weiter verwendet, auch wenn sie durch Tests widerlegt worden sind. So bleibt, auch wenn die Herleitung

[61] Allerdings sollen manche Fortschritte in Richtung Anwendbarkeit nicht geleugnet werden, etwa durch Saez (2001).
[62] Angesichts der Schwierigkeit, überzeugende empirische Erkenntnisse betreffend die Verzerrungseffekte von Steuern zu finden, mag man die dominante Rolle, die ihnen in der Theorie der optimalen Besteuerung - vor allem im Hinblick auf den Arbeitsmarkt - eingeräumt wird, nicht immer gerechtfertigt sehen. Jedenfalls ist sie, wie in der Einleitung dargelegt, erst im Laufe der letzten fünfzig Jahre so betont worden.

von Resultaten in Modellen selbst korrekt ist, ein weiter Spielraum für persönliche Intuition - und auch für Wertvorstellungen, welche wirtschaftspolitischen Schlussfolgerungen man für zulässig erachtet.

Das Denken in Modellen hat die deutschsprachige Tradition der Finanzwissenschaft weitgehend abgelöst, für die eine umfassende Abwägung vieler Aspekte kennzeichnend war (als ein Beispiel: Haller 1981). Als heutiger Leser wird man bei letzterer die Formulierung von Modellen vermissen, die eine genauere Abklärung der Bedingungen, unter denen manche Aussagen gelten, ermöglicht. Andererseits schränkt natürlich die Modellbildung die Vielfalt der behandelten Aspekte ein oder präformiert sogar durch die übliche Konzentration auf weitgehend akzeptierte Standardmodelle die Herangehensweise.

Am Ende einer Arbeit von Lucas, in der er davon spricht, dass die Anwendung der Resultate zur Abschaffung der Steuer auf Kapitaleinkommen einen beträchtlichen „free lunch“ ermöglicht, heißt es:

"As a practicing macroeconomist, I must say that I have greatly enjoyed this excursion into public finance. In my area, those of us who advocate structural modeling of aggregate behavior - accounting for observed behavior in terms of preferences and technology - remain very much on the defensive, accused of scientific utopianism and an excessive fascination with mathematical technique. How refreshing is it to spend some time in the company of a group of applied economists who simply take for granted the desirability of using (and extending) the powerful methods of dynamic general equilibrium theory to gain a deeper understanding of policy issues."
(Lucas 1990, p. 314.)

Würden Sie das als ein Kompliment an die Finanzwissenschaft verstehen?

Anhang 1

Differentiation der Lagrangefunktion

$$L = V(p_1 + \hat{\tau}, p_2 + \hat{\tau}_2, w, x) + \lambda(\sum_{i=1}^{2} \hat{\tau}_i c_i (p_1 + \hat{\tau}_1, p_2 + \hat{\tau}_2, w, x) - g)$$

nach $\hat{\tau}_k$, k = 1, 2 ergibt die Bedingungen erster Ordnung

$$\frac{\partial V}{\partial \hat{\tau}_k} + \lambda(c_k(\cdot) + \sum_{i=1}^{2} \hat{\tau}_i \frac{\partial c_i(\cdot)}{\partial \hat{\tau}_k}) = 0, \quad k = 1,2.$$

Mit $\gamma \equiv \partial V / \partial x$ als Grenznutzen des Einkommens folgt unter Verwendung von Roys Lemma:

$$(-\gamma + \lambda) c_k(\cdot) + \lambda \sum_{i=1}^{2} \hat{\tau}_i \frac{\partial c_i(\cdot)}{\partial \tau_k} \equiv 0, \qquad k = 1, 2,$$

bzw.

$$\sum_{i=1}^{2} \hat{\tau}_i \frac{\partial c_i(\cdot)}{\partial \hat{\tau}_k} = -\frac{\lambda - \gamma}{\lambda} c_k(\cdot), \qquad k = 1, 2. \qquad \text{(A1.1)}$$

Daraus ergibt sich mithilfe von $\partial c_i(\cdot) / \partial q_k = \partial c_i^c(\cdot) / \partial q_k - c_k(\cdot) \partial c_i(\cdot) / \partial x$ (Slutsky-Gleichung), wobei $c_i^c(\cdot)$ die kompensierte Nachfragefunktion nach Gut i bezeichnet und klarerweise $\partial c_i / \partial \hat{\tau}_k = \partial c_i / \partial q_k$ gilt:

$$\sum_{i=1}^{2} \hat{\tau}_i \frac{\partial c_i^c(\cdot)}{\partial \hat{\tau}_k} = -(\frac{\lambda - \gamma}{\lambda} - \sum_{i=1}^{2} \hat{\tau}_i \frac{\partial c_i(\cdot)}{\partial x}) c_k(\cdot), \quad k = 1,2.$$

Bezeichnet man den Klammerausdruck auf der rechten Seite, der unabhängig von k ist, mit θ[63] und verwendet die Symmetrie der kompensierten Preiseffekte, $\partial c_i^c(\cdot) / \partial \hat{\tau}_k = \partial c_k^c(\cdot) / \partial \hat{\tau}_i$ (siehe etwa Varian 1992, S. 123), so ergibt sich Gleichung (B.I.4) im Text.

[63] Man kann zeigen, dass θ positiv ist (falls g > 0). Siehe Atkinson und Stiglitz 1980, S. 373.

Zur Intuition von Samuelson gelangt man, wenn man den Effekt (kleiner) proportionaler Änderungen $\Delta\hat{\tau}_i = \beta\hat{\tau}_i$, i = 1, 2, mit gleichem Faktor β für alle Steuern, auf die kompensierte Nachfrage nach Gut k, ausgehend vom Optimum betrachtet. Offensichtlich gilt für die Änderung Δc_k^c (totales Differential)

$$\Delta c_k^c = \sum_{i=1}^{2} \Delta\hat{\tau}_i \frac{\partial c_k^c(\cdot)}{\partial \hat{\tau}_i} = \beta \sum_{i=1}^{2} \hat{\tau}_i \frac{\partial c_k^c(\cdot)}{\partial \hat{\tau}_i} = \beta\theta c_k^c(\cdot),$$

somit $\Delta c_k^c / c_k^c = \beta\theta$, wobei $\beta\theta$ eine Konstante, unabhängig von k, ist.

Schließlich betrachten wir noch den Fall, dass die Kreuzpreiseffekte null sind. Dann wird (A1.1) zu

$$\hat{\tau}_k \frac{\partial c_k(\cdot)}{\partial \hat{\tau}_k} = -\frac{\lambda - \gamma}{\lambda} c_k(\cdot),$$

wobei wir $\partial c_k / \partial \hat{\tau}_k < 0$ unterstellen, somit $(\lambda - \gamma)/\lambda > 0$. Mit ε_k als (Absolutbetrag der) Preiselastizität der Nachfrage für Gut k folgt

$$\frac{\hat{\tau}_k}{p_k + \hat{\tau}_k} = \frac{\lambda - \gamma}{\lambda} \frac{1}{\varepsilon_k}.$$

Dieser Zusammenhang beschreibt die Inverse-Elastizitäten-Regel, da der Quotient auf der linken Seite streng monoton steigend in $\hat{\tau}_k$ ist. (Offensichtlich lässt sich $\hat{\tau}_k /(p_k + \hat{\tau}_k)$ auch in der Form $\tau_k /(1 + \tau_k)$ schreiben, mit τ_k als Wertsteuersatz auf Gut k.)

Anhang 2

Mithilfe der Lagrangefunktion $L = \rho^1 v^1(z^1,x^1) + \rho^2 v^2(z^2,x^2) + \mu_1(v^1(z^1,x^1) - v^1(z^2,x^2)) +$ $\mu_2(v^2(z^2,x^2) - v^2(z^1,x^1)) - \nu(x^1 + x^2 - z^1 - z^2 + g - \hat{\tau}_1(c_1^1 + c_1^2) - \hat{\tau}_2(c_2^1 + c_2^1))$, wobei $\mu_1,\mu_2,\nu \geq 0$ die Lagrangevariablen zu (B.II.2), (B.II.3), (B.II.4)) bezeichnen, ergeben sich die Bedingungen erster Ordnung für den Fall einer inneren Lösung:

$$\rho^1\frac{\partial v^1}{\partial z^1} + \mu_1\frac{\partial v^1}{\partial z^1} - \mu_2\frac{\partial v^2}{\partial z^1} + \nu + \nu\hat{\tau}_1\frac{\partial c_1^1}{\partial z^1} + \nu\hat{\tau}_2\frac{\partial c_2^1}{\partial z^1} = 0\,, \tag{A2.1}$$

$$\rho^1\frac{\partial v^1}{\partial x^1} + \mu_1\frac{\partial v^1}{\partial x^1} - \mu_2\frac{\partial v^2}{\partial x^1} - \nu + \nu\hat{\tau}_1\frac{\partial c_1^1}{\partial x^1} + \nu\hat{\tau}_2\frac{\partial c_2^1}{\partial x^1} = 0\,, \tag{A2.2}$$

$$\rho^2\frac{\partial v^2}{\partial z^2} - \mu_1\frac{\partial v^1}{\partial z^2} + \mu_2\frac{\partial v^2}{\partial z^2} + \nu + \nu\hat{\tau}_1\frac{\partial c_1^2}{\partial z^2} + \nu\hat{\tau}_2\frac{\partial c_2^2}{\partial z^2} = 0\,, \tag{A2.3}$$

$$\rho^2\frac{\partial v^2}{\partial x^2} - \mu_1\frac{\partial v^1}{\partial x^2} + \mu_2\frac{\partial v^2}{\partial x^2} - \nu + \nu\hat{\tau}_1\frac{\partial c_1^2}{\partial x^2} + \nu\hat{\tau}_2\frac{\partial c_2^2}{\partial x^2} = 0\,. \tag{A2.4}$$

Wir betrachten den Fall $\hat{\tau}_1 = \hat{\tau}_2 = 0$. Man kann zeigen, dass strikte Konkavität von U und $\rho^1 \geq \rho^2$ (oder Konkavität von U und $\rho^1 > \rho^2$) implizieren, dass im Optimum der soziale Grenznutzen des Einkommens von Person eins größer ist als der von Person zwei, also $\rho^1\partial v^1/\partial x^1 > \rho^2\partial v^2/\partial x^2$ gilt (Brunner 1995, 1989, S. 190 ff). Außerdem ist im Optimum $\nu > 0$ (die Ressourcenbedingung bindet, wie man sich leicht überlegt). Daraus folgt, $\mu_2 > 0$. Nimmt man nämlich $\mu_2 = 0$ an, so ergibt sich aus (A2.2) und (A2.4)

$$\rho^1\frac{\partial v^1}{\partial x^1} - \rho^2\frac{\partial v^2}{\partial x^2} + \mu_1\frac{\partial v^1}{\partial x^1} = -\mu_1\frac{\partial v^1}{\partial x^2},$$

ein Widerspruch, weil die linke Seite positiv ist (beachte, dass $\mu_1 \geq 0$ und für den Grenznutzen des Einkommens $\partial v^1/\partial x^1 > 0$ gilt), die rechte Seite nicht positiv.

Somit ist die Selbtselektionsbeschränkung für Person zwei (B.II.3) im Optimum bindend. Es folgt aus (A2.1) und (A2.2)

$$\frac{-\partial v^1(z^1,x^1)/\partial z^1}{\partial v^1(z^1,x^1)/\partial x^1} = \frac{\nu - \mu_2\partial v^2(z^1,x^1)/\partial z^1}{\nu + \mu_2\partial v^2(z^1,x^1)/\partial x^1} \tag{A2.5}$$

und weil wegen AM der Anstieg der Indifferenzkurve $MRS^h_{zx} = (-\partial v^h / \partial z) / (\partial v^h / \partial x)$ an der Stelle (z^1,x^1) für Person zwei kleiner als für Person eins ist, impliziert (A2.5) $MRS^2_{zx}(z^1,x^1) < MRS^1_{zx}(z^1,x^1) < 1$, also insbesondere (B.III.1) im Text.[64]

Aus (A2.3) und (A2.4) folgt

$$\frac{-\partial v^2(z^2,x^2)/\partial z^2}{\partial v^2(z^2,x^2)/\partial x^2} = \frac{\nu - \mu_1 \partial v^1(z^2,x^2)/\partial z^2}{\nu + \mu_1 \partial v^1(z^2,x^2)/\partial x^2}, \tag{A2.6}$$

woraus zunächst zu ersehen ist, dass $\mu_1 > 0$ nicht zutreffen kann. Offensichtlich implizieren nämlich im Fall $\mu_1 > 0$ (nach einer analogen Überlegung wie oben) AM und (A2.6), dass an der Stelle (z^2,x^2) für den Anstieg der Indifferenzkurven die Relation $1 < MRS^2_{zx}(z^2,x^2) < MRS^1_{zx}(z^2,x^2)$ gilt, ein Widerspruch zu obigen Ungleichung (beachte dabei, dass im Fall $\mu_1 > 0$, zusätzlich zu $\mu_2 > 0$, beide Selbstselektionsbedingungen (B.II.2), (B.II.3) binden, was bedeutet, dass beide Personen das gleich z-x-Bündel bekommen, $z^1 = z^2$, $x^1 = x^2$).

Eine weitere Überlegung zeigt, dass $\mu_1 = 0$ in (A2.6) unmittelbar die Charakterisierung $MRS^2_{zx}(z^2,x^2) = 1$, also (B.III.2) im Text, ergibt.

[64] Wenn a/b > c/d für positive Zahlen a,b,c,d, so kann a/b = (ν + c)/(ν + d), für ν > 0, nur dann gelten, wenn a/b < 1.

Anhang 3

Wir verwenden die in Anhang 2 formulierte Lagrangefunktion und erhalten gemäß dem Envelope-Theorem

$$\frac{\partial S}{\partial \hat{\tau}_i} = \rho^1 \frac{\partial v^1}{\partial \hat{\tau}_i} + \rho^2 \frac{\partial v^2}{\partial \hat{\tau}_i} + \mu_1\left(\frac{\partial v^1}{\partial \hat{\tau}_i} - \frac{\partial v^1[2]}{\partial \hat{\tau}_i}\right) + \mu_2\left(\frac{\partial v^2}{\partial \hat{\tau}_i} - \frac{\partial v^2[1]}{\partial \hat{\tau}_i}\right) + \nu\left(c_i^1 + c_i^2 + \sum_{j=1}^{2} \hat{\tau}_j \left(\frac{\partial c_j^1}{\partial \hat{\tau}_i} + \frac{\partial c_j^2}{\partial \hat{\tau}_i}\right)\right),$$

$$i = 1, 2.$$

Dabei besagt der Ausdruck [2] bzw. [1], dass die partielle Ableitung von v^1 an der Stelle (z^2,x^2) bzw. von v^2 an der Stelle (z^1,x^1) zu bilden ist.

Aus der Umverteilungszielsetzung des Staates ergibt sich, wie in Anhang 2 erörtert, dass $\mu_1 = 0$, $\mu_2 > 0$ (die Selbstselektionsbedingung (B.II.3) bindet). Aus der Definition der in B.II. eingeführten Nutzenfunktion $v^h(z^h,x^h,p_1+\hat{\tau}_1,p_2+\hat{\tau}_2)$ und Roys Lemma folgt $\partial v^h / \partial \hat{\tau}_i = -c_i \partial v^h / \partial x^h$, so dass insgesamt an der Stelle $\hat{\tau}_1 = \hat{\tau}_2 = 0$ gilt

$$\left.\frac{\partial S}{\partial \hat{\tau}_i}\right|_{\tau_1=\tau_2=0} = -\rho^1 c_k^1 \frac{\partial v^1}{\partial x^1} - \rho^2 c_k^2 \frac{\partial v^2}{\partial x^2} + \nu(c_i^1 + c_i^2) - \mu_2 c_i^2 \frac{\partial v^2}{\partial x^2} + \mu_2 c_i^2[1] \frac{\partial v^2[1]}{\partial x^1}, \quad i = 1, 2.$$

Multipliziert man schließlich (A2.2) mit c_i^1 und (A2.4) mit c_i^2, so kann man νc_i^1 und νc_i^2 eliminieren und man erhält die Formel (B.III.3) im Text (beachte, dass $\mu_1 = \hat{\tau}_1 = \hat{\tau}_2 = 0$).

E. Literatur

Aiyagari, R. (1995), Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting, Journal of Political Economy 103, 1158-75.

Atkeson, A., V. V. Chari und P. J. Kehoe (1999), Taxing capital income: a bad idea, Federal Reserve Bank of Minneapolis Quarterly Review 23, 3-17.

Atkinson, A. B. (1973), How progressive should income tax be?, in: Essays in Modern Economics, M. Parkin und A. R. Nobay (Hrsg.), London: Longman.

Atkinson, A. B. (1977), Optimal taxation and the direct versus indirect tax controversy, Canadian Journal of Economics 10, 590-606.

Atkinson, A. B und A. Sandmo (1980), Welfare implications of the taxation of savings, Economic Journal 90, 529-549.

Atkinson, A. B. und J. Stiglitz (1976), The Design of tax structure: direct versus indirect taxation, Journal of Public Economics 6, 55-75.

Atkinson, A. B und J. Stiglitz (1980), Lectures on public economics, New York und London: McGraw-Hill.

Auerbach, A J. und J. R. Hines Jr. (2002), Taxation and economy efficiency, in: A. J. Auerbach und M. Feldstein (Hrsg.), Handbook of Public Economics, Vol 3, Amsterdam et al.: Elsevier.

Barro, Robert J. (1974), Are government bonds net wealth?, Journal of Political Economy 82, 1095-1117.

Benhabib, J. und A. Rustichini (1997), Optimal taxes without commitment, Journal of Economic Theory 77, 231-259.

Bernheim, D. (2002), Taxation and saving, in: A. J. Auerbach und M. Feldstein (Hrsg.), Handbook of Public Economics, Vol. 3, Amsterdam et al.: Elsevier.

Berliant, M. und M. Gouveia (1993), Equal sacrifice and incentive compatible income taxation, Journal of Public Economics 51, 219-240.

Boadway, R. und M. Keen (1993), Public Goods, Self-selection and optimal income taxation, International Economic Review 34, 463-478.

Boadway, R., Marchand M. und P. Pestieau (1994), Towards a theory of the direct-indirect tax mix, Journal of Public Economics 55, 71-88.

Boadway, R., Marchand M. und P. Pestieau (2000), Redistribution with unobservable bequests: a case for taxing capital income, Scandinavian Journal of Economics 102, 253-267.

Boskin, M. und E. Sheshinksi (1978), Optimal redistributive taxation when individual welfare depends upon relative income, Quarterly Journal of Economics 92, 589-601.

Brunner, Johann K. (1989), Theory of equitable taxation, Springer Verlag.

Brunner, Johann K. (1990), Optimale Pensionsversicherung im Overlapping-generations-Modell, Finanzarchiv 48, 467-484.

Brunner, Johann K. (1990), Utilitarian income taxation, Arbeitspapier des Instituts für Volkswirtschaftslehre Nr. 9003, Universität Linz.

Brunner, Johann K. (1995), A theorem on utilitarian redistribution, Social Choice and Welfare 12, 175-179.

Brunner, Johann K. (1997), Optimal taxation of income and bequests, Arbeitspapier des Instituts für Volkswirtschaftslehre Nr. 9722, Universität Linz.

Brunner, Johann K. (1999), Optimum taxation of income and consumption when individuals differ in abilities and endowments, mimeo, Universität Linz.

Brunner, Johann K. (2003), Optimum taxation of income from labour and capital in a dynamic two-person economy, Arbeitspapier des Instituts für Volkswirtschaftslehre Nr. 0309, Universität Linz.

Chamley, C. (1986), Optimal taxation of capital income in general equilibrium with infinite lives, Econometrica 54, 607-622.

Chamley, C. (2001), Capital income taxation, wealth distribution and borrowing constraints, Journal of Public Economics 79, 55-69.

Chari, V. V. und P. J. Kehoe (1999), Optimal fiscal and monetary policy, in: J. B. Taylor and M. Woodford (Hrsg.), Handbook of Macroeconomics, Vol. 1C, Amsterdam et al.: Elsevier.

Christiansen, V. (1981), Evaluation of public projects under optimal taxation, Review of Economic Studies 48, 447-457.

Cohen-Stuart, A. J. (1889), On progressive taxation, reprinted in: R. A. Musgrave and A. T. Peacock (Hrsg.), Classics in the theory of public finance, London et. al.: MacMillan 1967.

Correia, I. H. (1996), Should capital income be taxed in the steady state?, Journal of Public Economics 60, 147-151.

Corlett, W. J. und D. C. Hague (1953), Complementarity and the excess burden of taxation, Review of Economic Studies 21, 21-30.

Corneo, G. (2002), The efficient side of progressive income taxation, European Economic Review 46, 1359-1368.

Cremer, H., P. Pestieau und J.-Ch. Rochet (2001), Direct versus indirect taxation: the design of the tax structure revisted, International Economic Review 42, 781-799.

Daran, M. und M. Strawczynski (2000), Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates: comment, American Economic Review 90, 681-686.

Deaton, A. (1981), Optimal taxes and the structure of preferences, Econometrica 49, 1245-1260.

Diamond, P. A. (1965), National debt in a neoclassical growth model, American Economic Review 55, 1126-1150.

Diamond, Peter (1980), Income taxation with fixed hours of work, Journal of Public Economics 13, 101-110.

Diamond, Peter A. (1998), Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates, American Economic Review 88, 83-95.

Diamond, P. A. und J. A. Mirrlees (1971), Optimal taxation and public production I: production efficiency, American Economic Review 61, 8-27.

Eaton, J. und H. S. Rosen (1980), Taxation, human capital and uncertainty, American Economic Review. 70, 705-715.

Ebert, Udo (1992), A reexamination of the optimal nonlinear income tax, Journal of Public Economics 49, 47-73.

Edgeworth, F. Y. (1897), The pure theory of taxation, Economic Journal 7, reprinted in: R. A. Musgrave and Peacock A. T. (Hrsg.), Classics in the theory of public finance, London et al.: MacMillan 1967.

Edwards, J., Keen M. und M. Tuomala (1994), Income tax, commodity taxes and public good provision: a brief guide, Finanzarchiv 51, 472-487.

Erosa, A. und M. Gervais (2002), Optimal taxation in life-cycle economies, Journal of Economic Theory 105, 338-369.

Feldstein, M. (1973), On the optimal progressivity of the income tax, Journal of Public Economics 2, 357-376.

Feldstein, M. (1995), The effect of marginal tax rates on taxable income: a panel study of the 1986 tax reform act, Journal of Political Economy 103, 551-572.

Frisch, R. (1932), New methods of measuring marginal utility. in: Lederer, E. und J. Schumpeter (Hrsg.), Beiträge zur ökonomischen Theorie 3, Tübingen, J.C.B Mohr.

Goolsbee, A. (2000), What happens when you tax the rich? Evidence from executive compensation, Journal of Political Economy 108, 352-378.

Guesnerie, R. und J. Seade (1982), Nonlinear pricing in a finite economy, Journal of Public Economics 17, 157-179.

Gruber, J. und E. Saez (2002), The elasticity of taxable income: evidence and implications, Journal of Public Economics 84, 1-32.

Haller, H. (1981), Die Steuern, 3. Auflage, Tübingen: J.C.B Mohr.

Hellwig, M. F. (1986), The optimum linear income tax revisited, Journal of Public Economics 31, 163-179.

Homburg, Stefan (2000), Allgemeine Steuerlehre, 2. Auflage, Wiso-Kurzlehrbücher, Verlag Vahlen.

Homburg, Stefan (2001), The optimal income tax: Restatement and Extensions, Finanzarchiv 58, 363-395.

Homburg, Stefan (2002), Optimal marginal tax rates for low incomes: positive, negative, or zero?, Discussion Paper No 255, University of Hannover.

Ireland, N. J. (2001), Optimal income tax in the presence of status effects, Journal of Public Economics 81, 193-212.

Itsumi, Y. (1974), Distributional effects of linear income tax schedules, Review of Economic Studies 41, 371-381.

Jones, L., Manuelli, R. und P. Rossi u. a. (1997), On the optimal taxation of capital income, Journal of Economic Theory 73, 93-117.

Judd, K. L. (1985), Redistributive taxation in a simple perfect foresight model, Journal of Public Economics 28, 59-83.

Krusell, P. (2002) Time-consistent redistribution, European Economic Review 46, 755-769.

Lansing, K. J. (1999), Optimal redistributive capital taxation in a neoclassical growth model, Journal of Public Economics 73, 423-453.

Lucas, R. E. Jr. (1990), Supply-side economics: an analytical review, Oxford Economic Papers 42, 293-316.

Mirrlees, J. (1971), An exploration in the theory of optimal income taxation, Review of Economic Studies 38, 135-208.

Mirrlees, J. (1974), Notes on welfare economics, information, and uncertainty. In: Balch, M. S., D. L. McFadden and S.Y. Wu (eds.), Essays on economic behavior under uncertainty. Amsterdam: North-Holland, 243-258.

Munk, K. J. (1980), Optimal taxation with some non-taxable commodities, Review of Economic Studies 47, 755-765.

Naito, H. (1999), Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency, Journal of Public Economics 71, 165-188.

Ordover, J. A. und E. S. Phelps (1979), The concept of optimal taxation in the overlapping-generations model of capital and wealth, Journal of Public Economics 12, 1-26.

Oswald, A. (1983), Altruism, jealousy, and the theory of optimal non-linear taxation, Journal of Public Economics 20, 77-88.

Pestieau, P. M. (1974), Optimal taxation and discount rate for public investment in a growth setting, Journal of Public Economics 3, 217-235.

Ramsey, F. P. (1927), A contribution to the theory of taxation, Economic Journal 37, 47-61.

Rose, M., H. D. Henzel und W. Wiegard (1981), Optimale Finanzpolitik - Ein Lehr- und Arbeitsbuch, Stuttgart - New York: Gustav Fischer Verlag.

Sadka, E. (1976), On income distribution, incentive effects and optimal taxation, Review of Economic Studies 43, 261-267.

Saez, E. (2001), Using elasticities to derive optimal income tax rates, Review of Economic Studies 68, 205-229.

Saez, E. (2002a), Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses, Quarterly Journal of Economics 117, 1039-1073.

Saez, E. (2002b), The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes, Journal of Public Economics 83, 217-230.

Sandmo, A. (1993), Optimal redistribution when tastes differ, Finanzarchiv 50, 149-163.

Seade, J. K. (1977), On the shape of optimal tax schedules, Journal of Public Economics 7, 203-235.

Seade, J. K. (1982), On the sign of the optimum marginal income tax, Review of Economic Studies 49, 637-643.

Sheshinksi, E. (1972), The optimal linear income-tax, Review of Economic Studies 39, 297-302.

Stiglitz, Joseph E. (1982), Self-selection and pareto efficient taxation, Journal of Public Economics 17, 213-240.

Strawczynski, M. (1998), Social insurance and the optimum piecewise linear income tax, Journal of Public Economics 69, 371-388.

Tuomala, Matti (1990), Optimal income tax and redistribution, Oxford University Press.

Varian, H. R. (1980), Redistributive taxation as social insurance, Journal of Public Economics 14, 49-68.

Varian, H. R. (1992), Microeconomic analysis, 3rd edition, New York - London: W.W. Norton & Company.

Vogelsang, U. (2000), Optimal capital income taxation and redistribution, Finanzarchiv 57, 412-434.

Wane, W. (2001), The optimal income tax when poverty is a public "bad", Journal of Public Economics 82, 271-299.