

Sielaff, Christian

Working Paper

Steuerkomplexität und Arbeitsangebot: Eine experimentelle Analyse

Diskussionsbeiträge, No. 2011/13

Provided in Cooperation with:

Free University Berlin, School of Business & Economics

Suggested Citation: Sielaff, Christian (2011) : Steuerkomplexität und Arbeitsangebot: Eine experimentelle Analyse, Diskussionsbeiträge, No. 2011/13, Freie Universität Berlin, Fachbereich Wirtschaftswissenschaft, Berlin

This Version is available at:

<https://hdl.handle.net/10419/49297>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Steuerkomplexität und Arbeitsangebot - Eine experimentelle Analyse

Christian Sielaff

Fachbereich Wirtschaftswissenschaft

Diskussionsbeiträge

FACTS

2011/13

Steuerkomplexität und Arbeitsangebot - Eine experimentelle Analyse

Christian Sielaff*

August 2011

Abstract

Die vorliegende Arbeit untersucht den Einfluss von komplexen Steuertarifsystemen auf die Arbeits-Freizeit-Entscheidung von Individuen. Dazu wird ein wirtschaftswissenschaftliches Experiment mit 96 studentischen Teilnehmern durchgeführt. Es wird eine realitätsnahe Arbeits-Freizeit-Entscheidung unter Setzung tatsächlicher monetärer Anreize modelliert, um so den Einfluss verschieden komplexer Steuertarifsysteme auf das individuelle Arbeitsangebot der Experimentteilnehmer abzuleiten. Die Modellierung der verschiedenen Komplexitätsstufen erfolgt dabei durch das Zusammenspiel mehrerer Steuersätze, wie es auch in der Besteuerungspraxis zu finden ist. Es kann gezeigt werden, dass bei konstantem Netto- und Bruttolohn das Arbeitsangebot der Teilnehmer allein durch eine Steigerung des Komplexitätsgrades des Steuertarifsystems signifikant sinkt. Die Fragestellung ist sowohl für die theoretische Forschung als auch für zukünftige Steuerreformdebatten mit dem Ziel der Steuervereinfachung von Interesse.

Keywords

Steuerkomplexität · Arbeits-Freizeit-Entscheidung · Experimentelle Steuerforschung · Real-Effort-Experiment · Behavioral Taxation

JEL classification

C91 · H24 · H31

* Dipl.-Ök. Christian Sielaff, Lehrstuhl für Betriebswirtschaftliche Steuerlehre, Fachbereich Wirtschaftswissenschaft, Freie Universität Berlin, Email: christian.sielaff@fu-berlin.de

Inhaltsverzeichnis

1. Einleitung.....	4
2. Literaturüberblick.....	6
3. Spezifikation der Forschungsfrage und Experimentdesign	9
3.1 Spezifikation der Forschungsfrage.....	9
3.2 Experimentdesign.....	12
4. Ergebnisse und Diskussion	16
5. Zusammenfassung und Ausblick	33
Anhang.....	35
Literaturverzeichnis	39

Tabellenverzeichnis

Tabelle 1 - Steuertarifsysteme.....	14
Tabelle 2 - Gesamte Arbeitsleistung im Experiment.....	16
Tabelle 3 - Produktivitäten im Zeitablauf	17
Tabelle 4 - Arbeitsleistung nach Gruppen	19
Tabelle 5 - Paarweise Gruppenvergleiche	21
Tabelle 6 - Produktivitäten nach Zeitintervall und Steuertarifsystem	22
Tabelle 7 - Arbeitszeit und Arbeitsleistung in korrigierter Form	23
Tabelle 8 - Paarweise Gruppenvergleiche in korrigierter Form	24
Tabelle 9 - Fragen zum Experiment.....	26
Tabelle 10 - OLS Regression, Abhängige Variable: Arbeitszeit, Erweiterte Spezifikation	29
Tabelle 11 - OLS Regression, Abhängige Variable: Arbeitszeit, Basisspezifikation.....	30
Tabelle 12 - OLS Regression, Abhängige Variable: Arbeitsleistung, Erweiterte Spezifikation	31
Tabelle 13 - OLS Regression, Abhängige Variable: Arbeitsleistung, Basisspezifikation.....	32

Abbildungsverzeichnis

Abbildung 1 - Produktivitäten im Zeitablauf	18
Abbildung 2 - Arbeitszeit und Arbeitsleistung	20

1. Einleitung

Die Komplexität des deutschen Einkommensteuersystems steht immer wieder im Mittelpunkt zahlreicher politischer wie auch wissenschaftlicher Diskussionen. Thematisiert werden dabei insbesondere Fragen der Steuervereinfachung, welche regelmäßig von verschiedenen Seiten gefordert werden.¹ Zu welchen konkreten Verhaltensreaktionen auf Seiten der Steuerpflichtigen ein komplexes Steuertarifsystem führt, bleibt bei diesen Debatten allerdings häufig außer Acht. Der vorliegende Beitrag widmet sich deshalb dieser Frage. Es soll anhand einer experimentellen Untersuchung gezeigt werden, welchen Einfluss die Komplexität eines Steuertarifsystems auf Arbeitsangebotsentscheidungen eines Steuerpflichtigen ausübt.

Die Reaktion der Steuerpflichtigen auf Steuerkomplexität ist von zentraler Bedeutung, um über die Notwendigkeit und die Folgen einer Komplexitätsreduzierung innerhalb des deutschen Steuertarifsystems sinnvoll debattieren zu können, bzw. um über die Umsetzung von Steuervereinfachungen zu entscheiden. Die Durchführung eines Experiments erscheint dabei vor dem Hintergrund dieser Fragestellung als besonders geeignet, da durch ein Experiment ein realitätsnahes Umfeld geschaffen werden kann, welches erlaubt, das Verhalten und die Reaktionen von Steuerpflichtigen unter kontrollierten Bedingungen zu beobachten. Aus diesen experimentellen Ergebnissen können Prognosen über die Auswirkungen einer Steuervereinfachung abgeleitet werden. Insbesondere das Setzen tatsächlicher monetärer Anreize erlaubt eine angemessene Modellierung einer realen Arbeits-Freizeit-Entscheidung und eine Untersuchung des Einflusses einer tatsächlich zu zahlenden Steuerbelastung.

Die Untersuchung ist somit sowohl für die theoretische betriebswirtschaftliche Forschung, insbesondere für die steuerliche Verhaltensforschung (Behavioral Taxation), wie auch für die aktuelle politische Diskussion von großer Relevanz.

In der steuerökonomischen Modelltheorie wird im Allgemeinen unterstellt, dass Steuerpflichtige das jeweilige Steuertarifsystem vollständig verstehen und somit in der Lage sind, steuerliche Regelungen und Vorschriften im Rahmen ihrer Steuerplanungs- und Steuergestaltungsmaßnahmen

¹ Vgl. dazu bspw. die Ziele der Koalitionsparteien im Bundestag, Koalitionsvertrag zwischen CDU, CSU und FDP (2009) sowie den Entwurf eines Steuervereinfachungsgesetzes 2011, BT-Drucks. 17/5125. Ein aktueller Reformentwurf zur Erneuerung und Vereinfachung des Steuerrechts findet sich bei *Kirchhof* (2011). Dazu kritisch: *Wagner* (2011) sowie *Bach* (2011).

in rationaler Weise zu nutzen.² Aus diesen Modellen wird dann auf das tatsächliche Verhalten in der Realität sowie ggf. auf die damit verbundenen gesamtwirtschaftlichen Auswirkungen geschlossen. Sollte ein Großteil der Steuerpflichtigen aber aufgrund der Komplexität des Steuertarifsystems nicht in der Lage sein, steuerliche Sachverhalte in entsprechender Weise zu erkennen und so die damit verbundenen Planungs- und Gestaltungsmöglichkeiten zu nutzen, dann ist die Erklärungs- und Vorhersagekraft einer solchen Modelltheorie entsprechend niedriger. An dieser Stelle knüpft der vorliegende Beitrag an. Es soll empirisch untersucht werden, welche Verhaltensreaktionen Steuerpflichtige zeigen, wenn das zugrunde liegende Steuertarifsystem aufgrund der hohen Komplexität nicht verstanden wird. Diese Frage ist von großem theoretischen Interesse für die Steuerforschung, da die theoretischen Annahmen über die Wirkung der Besteuerung korrigiert werden müssen, wenn Steuerpflichtige nicht in der Lage sind, die Besteuerung korrekt in ihre ökonomischen Entscheidungen einfließen zu lassen, sondern ihr Verhalten vielmehr an der individuell wahrgenommenen Besteuerung und damit bspw. an Vereinfachungsregeln bzw. Heuristiken ausrichten.³

Weiterhin leistet diese Untersuchung einen wichtigen Beitrag zur steuerpolitischen Diskussion. Die Koalitionsparteien im Bundestag haben sich in ihrem Koalitionsvertrag darauf geeinigt, das Steuertarifsystem zu vereinfachen.⁴ Dieses Vorgehen steht im Gegensatz zur bisherigen Steuerpolitik, welche die Komplexität des Einkommensteuersystems in den letzten Jahren durch immer neue Sondertarifvorschriften wie die Einführung der Abgeltungsteuer oder die Thesaurierungsbegünstigung erhöht hat.⁵ Insgesamt existieren neben der Grundtarifvorschrift des § 32a EStG noch fünf weitere einkommensteuerliche Tarifnormen, die die Komplexität des Einkommensteuersystems wesentlich erhöhen.⁶ Um aus politischer Sicht zu entscheiden, welche Steuergesetzgebung vorzuziehen ist, sollten die politischen Entscheidungsträger auch die Auswirkungen steuerlicher Komplexität bzw. deren Vereinfachung bei den Zensiten berücksichtigen. Sollte neben der tatsächlichen Höhe der Besteuerung auch dem Komplexitätsgrad des Steuertarifsystems eine Bedeutung bei der Entscheidungsfindung seitens der Steuerpflichtigen zufallen, so muss dieser Effekt im

² So unterstellt bereits *Ramsey* (1927), dass Steuerpflichtige auf Steueränderungen in der gleichen Weise reagieren wie auf Preisänderungen. Diese Annahme wird auch in zahlreichen weiteren Beiträgen innerhalb der wirtschaftswissenschaftlichen Forschung übernommen, bspw. *Harberger* (1964), *Mirrlees* (1971), *Atkinson/Stiglitz* (1976).

³ Zur Anwendung von Heuristiken vgl. *Gigerenzer/Gaissmaier* (2011).

⁴ Vgl. Koalitionsvertrag zwischen CDU, CSU und FDP (2009).

⁵ Vgl. beispielsweise zur Abgeltungsteuer: *Hechtner/Hundsdoerfer* (2009), S. 41, oder zur Thesaurierungsbegünstigung: *Houben/Maiterth* (2008) sowie *Hechtner/Hundsdoerfer/Sielaff* (2011).

⁶ Vgl. *Hechtner* (2010), S. 27f.

Rahmen der Steuergesetzgebung berücksichtigt werden.⁷ Deshalb ist die hier untersuchte Fragestellung ebenfalls aus steuerpolitischer Sicht von Interesse.

Der vorliegende Beitrag gliedert sich wie folgt. Im zweiten Kapitel wird zunächst auf den aktuellen Stand der Literatur eingegangen. Kapitel 3 spezifiziert die untersuchte Forschungsfrage und beschreibt den Aufbau des Experiments. Kapitel 4 stellt die Ergebnisse des Experiments dar und diskutiert mögliche Ursachen und Implikationen. Kapitel 5 schließt diesen Beitrag mit einem zusammenfassenden Fazit sowie einem Ausblick auf zukünftigen Forschungsbedarf und geht auf Grenzen der Untersuchung ein.

2. Literaturüberblick

Die vorliegende Untersuchung knüpft insbesondere an die bisherige Literatur zu Wahrnehmung und Wirkung von komplexen Steuertarifsystemen an. *Cogden et al.* (2009) führen an, dass Steuerpflichtige nicht auf tatsächliche Steuersätze reagieren, sondern auf die individuell wahrgenommenen. Da gerade mit steigender Komplexität eines Steuertarifsystems der Spielraum für solche „Steuersatzinterpretationen“ steigt, scheint es naheliegend zu sein, diesem „interpretierten“ Steuersatz mehr Beachtung bei der Untersuchung von Steuerwirkungen zu schenken. Bisher wurde der Einfluss von Steuern auf ökonomische Entscheidungen in verschiedenen Studien untersucht, welche sich hinsichtlich der Modellierung des Steuertarifsystems, der genauen Fragestellung und der angewandten Methode stark unterscheiden.

So untersuchen *Gamage, Hayashi* und *Nakamura* (2010) im Rahmen eines Experiments den Einfluss der Präsentation des Nettoeinkommens auf das Arbeitsangebot. Bei diesem Experiment werden Probanden gebeten, sich nach Vorgabe eines bestimmten Nettolohns zu entscheiden, ob diese innerhalb eines bestimmten Zeitintervalls arbeiten wollen oder in dieser Zeit lieber einer vorgegebenen Freizeitbeschäftigung nachgehen wollen. Sie kommen zu dem Ergebnis, dass die Probanden eine höhere Arbeitsbereitschaft zeigen, wenn ihnen ein hoher Bruttolohn abzüglich einer

⁷ Der Komplexitätsgrad eines Steuertarifsystems könnte bspw. dazu führen, dass die Zensiten ihre Steuerbelastung anders wahrnehmen, als diese tatsächlich ist. Kleine Einzelsteuern könnten somit durch den Gesetzgeber im „Steuerwirrwarr“ versteckt werden und zu einer Differenz zwischen tatsächlicher und wahrgenommener Belastung führen, was wiederum den Steuerwiderstand beeinflussen kann. Zum subjektiven Belastungsgefühl vgl. *Schmölders* (1960), S. 80f. Vergleichbare Ansätze finden sich in der wirtschaftswissenschaftlichen Literatur bereits bei der Diskussion um die „Unmerklichkeit“ der Besteuerung. Hierbei werden die Vorteile einer indirekten Besteuerung im Vergleich zu einer direkten Besteuerung thematisiert, wobei ebenfalls die Senkung des Belastungsgefühl durch die Unmerklichkeit indirekter Steuern herausgestellt wird. Vgl. *Keller* (1968), S. 110; *Neumark* (1970), S. 37f.; *Tipke/Lang* (2010), § 8, Rn. 11. Dieses Vorgehen kann vom Grundsatz her auch auf Komplexitätswirkungen übertragen werden.

Steuer gezahlt wird anstatt eines niedrigen Nettolohns zzgl. einer Bonuszahlung. Hinsichtlich der Präsentation bzw. Darstellung von Steuern zeigen weiterhin *Chetty, Looney und Kroft* (2009) in einem Experiment, dass Steuern nur dann „richtig“ verarbeitet werden, wenn diese für den Steuerpflichtigen überhaupt erkennbar („salient“) sind. Auch das dient als Anknüpfungspunkt der vorliegenden Untersuchung, da die über mehrere Steuersätze modellierte Komplexität ebenfalls einen „salience effect“ darstellen kann. So werden die Teilnehmer ggf. durch eine der Komplexität geschuldeten längeren Beschreibung des Steuertarifsystems eher für die Steuerzahlung sensibilisiert.

Blumkin, Ruffle und Ganun (2008) modellieren im Rahmen eines Experiments ebenfalls eine Arbeits-Freizeit-Entscheidung, wobei die Probanden das verdiente Geld für reale Güter ausgeben müssen. Sie untersuchen dabei, welchen unterschiedlichen Einfluss eine Einkommensteuer und eine Konsumsteuer auf das Arbeitsangebot haben und kommen dabei zu dem Ergebnis, dass eine Einkommensteuer die Arbeitsbereitschaft stärker senkt als eine äquivalente Konsumsteuer.

Sillamaa (1999) untersucht im Rahmen einer Replikationsstudie zu *Swenson* (1988) den Einfluss verschiedener linearer Steuersätze auf das Arbeitsangebot und bestätigt das aus der Theorie bekannte Ergebnis, dass das Arbeitsangebot mit steigenden Steuersätzen und damit sinkendem Nettolohn sinkt.

Fochmann et al. (2010) beschäftigen sich mit den Auswirkungen eines Bruttolohneffekts bei der Arbeits-Freizeit-Entscheidung. Sie zeigen, dass das Arbeitsangebot von Probanden signifikant höher ist, wenn bei gleichem Nettolohn der Bruttolohn, aber dementsprechend auch der geltende Steuersatz, höher ist. Diesen Effekt führen *Fochmann et al.* (2008) auf eine „gross-wage illusion“ zurück.

Mit dem Einfluss der Komplexität auf die Wahrnehmung von Steuern beschäftigen sich *Chetty und Saez* (2009). Diese zeigen am Beispiel des Earned Income Tax Credit (EITC) in den USA, dass ein zusätzliches Informationsangebot zu Verhaltensanpassungen bei den Steuerpflichtigen führt, und dass der EITC zu kompliziert ist, um von den Steuerpflichtigen richtig wahrgenommen zu werden.

De Bartolome (1995) untersucht ebenfalls im Rahmen eines Experiments den Einfluss von Steuerkomplexität, indem er die Komplexität durch die unterschiedliche Präsentation des gleichen Steuertarifsystems variiert. *De Bartolome* (1995) kann zeigen, dass Steuern unabhängig von der Präsentationsform von den Probanden berücksichtigt werden, diese allerdings dazu neigen, den Durchschnittssteuersatz anstelle des Grenzsteuersatzes zu benutzen, wenn der Grenzsteuersatz nicht explizit bei der Darstellung des Steuertarifsystems ausgewiesen wird. Ein ähnliches Vorgehen findet sich bei *Rupert und Wright* (1998). Diese kommen zu dem Ergebnis, dass Steuerpflichtige

ihre Investitionsentscheidungen umso besser treffen, je transparenter ihnen der entsprechende Grenzsteuersatz präsentiert wird.

Ein weiteres Experiment zur Steuerkomplexität findet sich bei *Rupert et al.* (2003). Diese modellieren die Komplexität des Steuertarifsystems durch Abzüge von der Bemessungsgrundlage, welche dazu führen, dass die Berechnung des effektiven marginalen Steuersatzes komplexer wird. Bei diesem Experiment kann gezeigt werden, dass die Probanden seltener eine optimale Investitionsentscheidung treffen, wenn die Komplexität des Steuertarifsystems hoch ist. *Boylan* und *Frischmann* (2006) zeigen, ebenfalls anhand eines Experiments, dass ein komplexes Steuertarifsystem zu einer Beeinträchtigung der Markteffizienz führen kann.

Insbesondere die Wahrnehmung und Einschätzung von Grenzsteuersätzen wurde in der Literatur häufig behandelt. So finden sich dazu diverse Untersuchungen etwa bei *Gensemer et al.* (1965), *Lewis* (1978), *Fujii* und *Hawley* (1988), *Rupert* und *Fisher* (1995) sowie im Rahmen einer Conjoint-Analyse bei *Hundsdoerfer* und *Sichtmann* (2007). Diese Studien kommen größtenteils zu dem Schluss, dass Steuerpflichtige nicht in der Lage sind, ihren Grenzsteuersatz richtig zu berechnen und diesen somit für ökonomische Entscheidungen zu nutzen. Deshalb zeigen auch diese Untersuchungen, welche Folgen ein komplexes Steuertarifsystem und die damit verbundenen Fehleinschätzungen haben.

Ein anderes Vorgehen wählen *Blaufus* und *Ortlieb* (2009), welche zur Untersuchung von steuerlichen Komplexitätswirkungen auf die Methode der Conjoint-Analyse zurückgreifen. Steuerkomplexität wird hier anhand der Zeit, welche Steuerpflichtige brauchen, um ein Steuertarifsystem zu verstehen, gemessen. Die Steuertarifsysteme selbst unterscheiden sich dabei anhand der Länge der den Probanden vorgelegten Beschreibungen sowie durch die Anzahl der darin verwendeten Fachtermini. *Blaufus* und *Ortlieb* (2009) zeigen, dass mit steigender Komplexität immer weniger Steuerpflichtige die Steuerzahlungen bei ihren Entscheidungen berücksichtigen.

Das Verhältnis von Steuerkomplexität und Steuerehrlichkeit wurde im Rahmen eines Experiments von *Bizer* (2004, 2008) untersucht. *Bizer* zeigt, welchen Einfluss Unsicherheiten über die Höhe des zu versteuernden Einkommens auf die Bereitschaft zur Steuerhinterziehung ausüben. Er kommt zu dem Ergebnis, dass Komplexität Steuerehrlichkeit zerstört und Steuerpflichtige mehr Steuern hinterziehen, wenn die Höhe ihres zu versteuernden Einkommens unsicher ist. Ähnliche experimentelle Untersuchungen zum Verhältnis von komplexitätsbedingter Unsicherheit über die Steuerzahlung und Steuerehrlichkeit finden sich auch bei *Beck, Davis* und *Jung* (1991) sowie *Alm, Jackson* und *McKee* (1992).

Die hier durchgeführte Untersuchung baut auf den bisherigen experimentellen Studien auf und erweitert diese wie folgt. Zum einen wird die Arbeits-Freizeit-Entscheidung anders als in der bisherigen Literatur bekannt modelliert. Die oben aufgeführten Untersuchungen, mit Ausnahme von *Fochmann et al.* (2010), modellieren die Freizeitalternative bspw. durch das Hören von Musik, das Spielen von Computerspielen oder den Konsum von Freizeitgütern, was nur im Ansatz eine realistische Freizeitalternative darstellt. In der vorliegenden Untersuchung ist diese Entscheidung so ausgestaltet, dass die Probanden die entsprechende Arbeitstätigkeit jederzeit beenden und nach Hause gehen können.⁸ Dies entspricht eher der Entscheidungssituation in der Realität.

Wie oben bereits beschrieben, unterscheidet sich dieser Beitrag zum anderen durch die Modellierung der Steuersatzkomplexität und klammert dabei die bereits mehrfach untersuchten Effekte durch die Verwechslung von Grenz- und Durchschnittssteuersätzen aus. Der Beitrag widmet sich somit möglichen Komplexitätseffekten durch die Interaktion verschiedener linearer Steuertarife.

Außerdem werden tatsächliche Einkommenseffekte in diesem Experiment ausgeblendet. Im Gegensatz zu *Fochmann et al.* (2010) werden alle Probanden mit dem gleichen Bruttolohn konfrontiert, was eine „*gross-wage illusion*“ ausschließt. Auch der Einfluss des Nettolohns wird ausgeblendet, da sich die einzelnen Teilnehmergruppen nur hinsichtlich der Steuersatzkomplexität unterscheiden, der effektive Steuersatz sowie dementsprechend Brutto- und Nettolohn aber gleich sind. Insoweit ergänzt diese Untersuchung die bisherige Forschung durch eine noch nicht in dieser Form untersuchte Fragestellung, eine realitätsnahe Modellierung des Experimentrahmens und eine adäquate Modellierung der Steuerkomplexität.

3. Spezifikation der Forschungsfrage und Experimentdesign

3.1 Spezifikation der Forschungsfrage

Im Rahmen des Experiments werden die Teilnehmer mit einer realitätsnahen Arbeits-Freizeit-Entscheidung konfrontiert. Alle Teilnehmer erhalten während des Experiments den gleichen Bruttolohn für ihre Arbeitsleistung, welcher noch der Besteuerung unterworfen werden muss. Der jeweilige effektive Steuersatz und somit ebenfalls der Nettolohn sind gleich. Nur die Tarifkomplexität des zugrunde liegenden Steuersystems unterscheidet sich in vier verschiedenen Komplexitätsstufen. Die

⁸ *Fochmann et al.* (2010) wählen ein vergleichbares Vorgehen, stellen dabei aber eine andere Fragestellung in den Mittelpunkt.

Komplexität des Steuertarifsystems wird hier durch das Zusammenspiel verschiedener Steuersätze modelliert, um so die Frage zu beantworten, wie die Kombination aus mehreren Steuersätzen (und damit die Steuersatzkomplexität) die Wahrnehmung der Besteuerung bei den Steuerpflichtigen beeinflusst. Folglich stehen bei dieser Untersuchung Steuertarifeffekte im Mittelpunkt.⁹ Bemessungsgrundlageneffekte und Zeiteffekte der Besteuerung werden dabei nicht untersucht.¹⁰ Es wird gerade dieses Vorgehen gewählt, da sich hierfür zahlreiche Beispiele im nationalen¹¹ sowie in internationalen¹² Steuertarifsystemen finden lassen, die nahezu alle Steuerpflichtigen betreffen.

Die Modellierung eines progressiven Tarifverlaufs wird dabei in dieser Untersuchung bewusst ausgeklammert, da es das Ziel dieses Beitrags ist, die steuerlichen Komplexitätswirkungen durch das Zusammenspiel verschiedener Steuersätze zu untersuchen. Dieses praxisnahe Beispiel steuerlicher Tarifkomplexität soll nicht durch Fragen und Probleme der Progression weiter ausgeweitet werden.¹³ Auch im Hinblick auf die kognitiven Fähigkeiten der Teilnehmer erscheint dieses Vorgehen gerechtfertigt, da das Berechnen progressiver Steuersätze eine für diese Fragestellung unverhältnismäßige und auch nicht zielführende Erweiterung der Komplexität darstellen würde.

Es stellt sich folglich die Frage, wie Steuerpflichtige reagieren, wenn sie mit einem komplexen Steuersatzsystem konfrontiert werden. Es ist davon auszugehen, dass mit steigender Komplexität weniger Steuerpflichtige in der Lage sind, ihren effektiven Steuersatz zu berechnen, und folglich mehr Steuerpflichtige auf Heuristiken und persönliche Einschätzungen bzw. Vereinfachungen

⁹ Zur Typisierung steuerlicher Effekte in Zeiteffekte, Bemessungsgrundlageneffekte und Steuertarifeffekte vgl. *Wagner* (1984), S. 211f. sowie *Wagner* (1999), S. 478f.

¹⁰ *Blaufus et al.* (2009) zeigen bereits empirisch, dass Steuertarifeffekte von Steuerpflichtigen meist überschätzt werden, wohingegen Bemessungsgrundlageneffekte eher unterschätzt werden. Auch vor diesem Hintergrund scheint die Modellierung in diesem Experiment über den Steuertarif angemessen, um so die Teilnehmer für die Beachtung der Steuer zu sensibilisieren.

¹¹ Im nationalen Steuertarifsystem ist hierbei das Zusammenwirken von Einkommensteuer und Solidaritätszuschlag sowie Kirchensteuer zu nennen. Auch ein Vergleich mit dem Zusammenspiel von Einkommensteuer und Gewerbesteuer sowie deren Anrechnung nach § 35 EStG liegt dabei nahe. Diese Form der Steuersatzkomplexität wird in politischen Debatten, neben Vereinfachungsvorschlägen, welche allein die Bemessungsgrundlage betreffen, häufig aufgegriffen. So wurde die Selbstabzugsfähigkeit der Gewerbesteuer mit der Unternehmensteuerreform 2008 abgeschafft, was deren Berechnung stark vereinfachte. Weiterhin wird seit langem über eine generelle Abschaffung bzw. Reform der Gewerbesteuer und deren Integration in die Einkommen- bzw. Körperschaftsteuer diskutiert. Vgl. *Maiterth* (2003). Auch das von der FDP vorgeschlagene Stufenmodell zielt zu einem großen Teil auf Steuervereinfachungen durch Vereinfachung des Steuersatzes ab. Vgl. *FDP* (2010), S. 6.

¹² Im internationalen Kontext könnte es sich hier bspw. um das Zusammenspiel zwischen „local taxes“ und „federal taxes“ in den USA handeln, oder aber auch um das Zusammenspiel inländischer und ausländischer Steuersätze bei der Besteuerung internationaler Tätigkeiten, bei welchen es zur Anwendung des Anrechnungs- oder Abzugsverfahrens kommt.

¹³ Durch die Annahme linearer Steuersätze entsprechen sich Grenz-, Durchschnitts- und Differenzsteuersatz. Verwechslungen dieser Steuersätze seitens der Steuerpflichtigen wurden in der Literatur bereits zahlreich untersucht (vgl. Literaturüberblick) und sollen in diesem Beitrag nicht thematisiert und untersucht werden. Aus diesem Grund wird von progressiven Steuertarifen abgesehen.

zurückgreifen. Dabei ist es besonders von Interesse, ob die von den Steuerpflichtigen angewandten Heuristiken eher zu einer Über- oder Unterschätzung der tatsächlichen Steuerbelastung führen, was im Rahmen des Experiments die Arbeits-Freizeit-Entscheidung signifikant beeinflussen sollte.

In einer rationalen Welt, in der die Steuerpflichtigen die zu leistenden Steuerzahlungen richtig kalkulieren und antizipieren, sollte es keine signifikanten Unterschiede zwischen den verschiedenen Gruppen in diesem Experiment hinsichtlich des Arbeitsangebotes geben. Da die Berechnung des Nettolohns aber abhängig vom individuellen Steuertarifsystem mit unterschiedlichem Aufwand der Steuerberechnung verbunden ist, ist zu erwarten, dass sich insbesondere die Teilnehmer, welche mit einem komplexen Steuertarifsystem konfrontiert sind, eher eine Heuristik oder Vereinfachungsregel anwenden. Diese Teilnehmer werden dann ihr Arbeitsangebot anhand eines aus ihrer Sicht unsicheren Nettolohns wählen.

Die Forschungsfrage lässt sich somit weiter spezifizieren, indem untersucht wird, ob dieses Vorgehen eher zu einer Ausweitung oder Einschränkung des Arbeitsangebotes führt, was aus einer signifikanten Über- oder Unterschätzung des Steuersatzes und damit des Nettolohns resultieren kann.¹⁴ Die Aufdeckung der Richtung sowie die Quantifizierung dieses Verhaltens ist das Ziel dieses Beitrags. Außerdem werden Anhaltspunkte für die Grundlagen der individuell angewandten Heuristik herausgearbeitet. So könnte z. B. ein Ankereffekt¹⁵ vorliegen, der dazu führt, dass sich die Teilnehmer aus Vereinfachungsgründen an dem höchsten oder zuerst genannten Steuersatz orientieren und die sonstigen Steuern vernachlässigen. Auch eine vereinfachte Orientierung an der Summe der einzelnen Steuersätze wäre möglich. Bei der Modellierung der Steuertarifsysteme wurde deshalb

¹⁴ Die Wirkungsrichtung, die eine Über- oder Unterschätzung des Steuersatzes auf das Arbeitsangebot hat, kann im Vorfeld nicht eindeutig bestimmt werden. Bei den Teilnehmern kann eine Fehleinschätzung der zu zahlenden Steuerlast und damit des Nettolohns sowohl einen Einkommens- wie auch einen Substitutionseffekt auslösen. Wird bspw. der Steuersatz zu hoch und damit der Nettolohn zu niedrig eingeschätzt, kann der Einkommenseffekt zu einer Verringerung (da sich die Arbeit weniger lohnt) oder Ausweitung des Arbeitsangebotes (bspw. wenn ein bestimmter Mindestnettoloohn erzielt werden soll) führen. Der Substitutionseffekt führt dagegen immer zu einer Einschränkung des Arbeitsangebotes, da mit sinkendem Lohn Arbeitszeit durch Freizeit ersetzt wird. Ob der Einkommenseffekt und der Substitutionseffekt in die gleiche Richtung wirken oder ob der Substitutionseffekt vom Einkommenseffekt (über-)kompensiert wird hängt vom individuellen Verhalten der Teilnehmer ab. Zur Theorie vgl. bereits *Hicks* (1935). Empirische Überprüfungen der Arbeitsangebotsreaktion auf Lohnänderungen finden sich bspw. bei *Break* (1957), *Hausman* (1981), *Killingsworth* (1983), *Kimmel* und *Kniesner* (1998) sowie *Ziliak* und *Kniesner* (1999). Ein Überblick findet sich bei *Keane* (2010). Insgesamt deuten die bisherigen Untersuchungen eher auf einen positiven Zusammenhang zwischen Nettolohn und Arbeitsangebot hin, so dass auch hier die Vermutung einer Arbeitsangebotsreduktion (-ausweitung) bei (Unter-) Überschätzung der Steuerbelastung und somit (Über-) Unterschätzung des Nettolohns als Ausgangshypothese übernommen werden kann.

¹⁵ Ein Ankereffekt (*Tversky* und *Kahnemann* (1974)) bedeutet, dass sich die Probanden in der unsicheren Situation einen Startpunkt suchen, welchen diese dann beim Treffen ihrer Entscheidung zu stark gewichten. In dieser Untersuchung könnte der erste und höchste Steuersatz (Einkommensteuersatz) einen solchen Anker darstellen. Weitere empirische Überprüfungen eines Ankereffekts bei verschiedenen Fragestellungen finden sich bspw. bei *McCaffery* und *Baron* (2003), *Epley* und *Gilovich* (2006) sowie bei *Blaufus et al.* (2009).

darauf geachtet, dass ein solches Vorgehen der Teilnehmer untersucht werden kann. Der am Ende des Experiments vorgelegte Fragebogen enthält weiterhin spezifische Fragen, die Aufschluss darüber geben, inwieweit das Steuertarifsystem überhaupt Beachtung bei der Arbeitsangebotsentscheidung gefunden hat und ob der effektive Steuersatz vor Beginn der Arbeitstätigkeit berechnet wurde.

3.2 Experimentdesign

Das Experiment wurde in fünf Durchgängen von April 2011 bis Juni 2011 in den Räumen der Freien Universität Berlin durchgeführt. Dabei nahmen insgesamt 96 Probanden an dem Experiment teil. Bei den Probanden handelt es sich überwiegend um Studenten im zweiten Semester des Bachelorstudiengangs Betriebswirtschaftslehre.¹⁶ Auch Studenten anderer Fachbereiche waren unter den Teilnehmern vertreten. Da keine signifikanten Unterschiede zwischen den einzelnen Durchläufen sowie zwischen den Studenten der Betriebswirtschaftslehre und den Studenten anderer Fachbereiche existieren, werden diese Gruppen im folgenden zusammengefasst.¹⁷

Das Experiment wurde in den PC-Pools des Fachbereichs Wirtschaftswissenschaft durchgeführt. Die Teilnehmer wurden nach dem Zufallsprinzip auf die zur Verfügung stehenden Personal-Computer verteilt.¹⁸ Die jeweiligen Arbeitsplätze der Teilnehmer waren durch einen Sichtschutz voneinander abgetrennt, außerdem war die Kommunikation zwischen den Teilnehmern für die gesamte Dauer des Experiments untersagt. Bei keinem der Durchläufe waren mehr als 25 Teilnehmer gleichzeitig anwesend. Es standen bei jedem Durchlauf mehrere Experimentleiter zur Verfügung, welche den Ablauf des Experiments überwachten und auftretende Fragen beantworteten.

Zu Beginn des Experiments erhielten die Teilnehmer eine Anleitung ausgehändigt, die alle notwendigen Informationen zum Ablauf des Experiments beinhaltet.¹⁹ Im Anschluss daran wurde den Teilnehmern ausreichend Zeit gegeben, diese Anleitung sorgfältig durchzulesen und auftretende

¹⁶ Der Kritikpunkt, dass Studenten als Experimentteilnehmer ungeeignet bzw. nicht repräsentativ sind, wird in der Literatur bestritten. Vgl. *Plott* (1987), *Alm* (2010).

¹⁷ Zur Überprüfung wurde die Verteilung relevanter Variablen zwischen den einzelnen Durchläufen sowie den entsprechenden Studiengängen der Teilnehmer mit Hilfe von Kruskal-Wallis-Tests und Mann-Whitney-U-Tests überprüft. Es konnten keinerlei signifikante Unterschiede zwischen den Kategorien festgestellt werden. Des Weiteren wurde die Zugehörigkeit zum Studiengang Betriebswirtschaftslehre oder einem sonstigen Studiengang in die durchgeführte Regressionsanalyse aufgenommen, wobei sich allerdings auch dort kein signifikanter Einfluss nachweisen lässt.

¹⁸ Es gab keine signifikanten Unterschiede bei dieser Verteilung hinsichtlich der demografischen Merkmale: Alter, Geschlecht, Einkommen und Schulbildung bzw. Studium.

¹⁹ Die Anleitung findet sich im Anhang.

Fragen zu stellen. Die Teilnehmer wurden mit einer realen Arbeits-Freizeit-Entscheidung konfrontiert. Sie bekamen die Aufgabe, Lösungsbögen einer Multiple-Choice-Klausur zu Korrekturzwecken in eine vorgefertigte Exceldatei zu übernehmen. Jeder Lösungsbogen bestand aus 40 Fragen mit jeweils 6 Antwortmöglichkeiten, von denen jeweils eine Antwortmöglichkeit angekreuzt war. Bei der Wahl dieser Form einer Arbeitsaufgabe war es zum einen wichtig, dass die Teilnehmer das Gefühl hatten, eine realitätsnahe Aufgabe auszuführen und zum anderen sollte eine Aufgabe gewählt werden, welche keine Vorkenntnisse oder besondere Fähigkeiten erfordert. Für jeden fehlerfrei digitalisierten Bogen erhielten die Teilnehmer einen Bruttolohn in Höhe von 0,30 Euro, was ihnen ebenfalls in der Anleitung mitgeteilt wurde. Als Freizeitalternative zu dieser Arbeitsaufgabe wurden die Teilnehmer darüber informiert, dass diese das Experiment jederzeit beenden und nach Hause gehen können. Eine Zeitbeschränkung gab es nicht, jedem Teilnehmer stand es somit frei, so lange und soviel zu arbeiten wie er wollte.²⁰ So konnte eine realitätsnahe Entscheidungssituation geschaffen werden. Bei jedem Durchgang waren speziell instruierte zusätzliche Teilnehmer anwesend, die die Aufgabe hatten, nach 30 Minuten das Experiment zu beenden. Dadurch sollte die ggf. bestehende Hürde zur frühzeitigen Beendigung des Experiments beseitigt werden.

Der Bruttoarbeitslohn musste von den Teilnehmern versteuert werden. Dazu wurde den Teilnehmern im Anschluss an die Instruktionen noch eine Steuersystemkarte ausgehändigt, welche das für sie geltende Steuertarifsystem beinhaltet. Eine Auskunft über den Nettolohn selbst wurde den Teilnehmern nicht gegeben. Die Steuersystemkarten wurden jeweils zufällig auf die Teilnehmer verteilt, wobei darauf geachtet wurde, dass bei jedem Durchlauf, soweit möglich, jedem Steuersystem gleich viele Teilnehmer zugeordnet wurden.²¹ Dadurch konnte im Anschluss an das Experiment überprüft werden, ob mögliche Effekte auf durchlaufspezifische Unterschiede zurückzuführen sind.

Es gab insgesamt vier verschiedene Steuersystemkarten. Dabei unterschieden sich die Steuersysteme hinsichtlich ihrer Komplexität in der Form, dass jeweils eine unterschiedliche Anzahl an Steuersätzen zu beachten waren. Diese einzelnen Steuersätze mussten zur Berechnung des Gesamtsteuersatzes teilweise additiv oder multiplikativ verknüpft werden. Bei einer der Varianten (Steuersystem D) war eine Steuerzahlung zudem noch abzugsfähig von der Bemessungsgrundlage einer

²⁰ Das Experiment wurde ohne vorherige Ankündigung jeweils nach ca. 3 Stunden beendet, was über alle Durchläufe nur 5 Teilnehmer betraf. Alle anderen Teilnehmer haben das Experiment schon vorher, ohne Kenntnis des geplanten Abbruchzeitpunkts, beendet.

²¹ Die Steuertarifsystemkarten finden sich im Anhang.

weiteren Steuer. Die Besonderheit bei diesem Vorgehen bestand darin, dass der effektive Steuersatz bei allen Varianten gleich war und konstant 60% betrug.²² Lediglich die Darstellung bzw. der Komplexitätsgrad variierte zwischen den Steuertarifsystemen, was eine Berechnung des effektiven Steuersatzes an einen unterschiedlichen Aufwand bzw. unterschiedliche Kenntnisse knüpfte.

Variante	Einkommensteuer	Zusatzsteuer auf das Einkommen	Zuschlagsteuer auf die Einkommensteuer	Landessteuer auf das Einkommen	Gemeindesteuer mit Abzugsfähigkeit bei der Einkommensteuer	Effektiver Steuersatz	Komplexitätsmodellierung
A	60%	---	---	---	---	60,0%	Einfacher Steuersatz
B	40%	20%	---	---	---	60,0%	Additive Verknüpfung
C	40%	---	20%	12%	---	60,0%	Additive und multiplikative Verknüpfung
D	20%	---	20%	20%	20%	59,2%	Additive und multiplikative Verknüpfung sowie Abzugsfähigkeit

Tabelle 1 - Steuertarifsysteme

Während bei Variante A lediglich ein pauschaler Steuersatz in Höhe von 60% erhoben wurde, musste dieser bei Variante B über eine einfache additive Verknüpfung ermittelt werden ($40\% + 20\% = 60\%$). Für die Ermittlung des effektiven Gesamtsteuersatzes war bei Variante C dann zusätzlich eine multiplikative Rechenoperation notwendig. Hier ergab sich der Gesamtsteuersatz als $40\% \cdot (1 + 20\%) + 12\% = 60\%$. Bei Variante D wurde dann das Rechenverfahren nochmals in der Form erweitert, dass zusätzlich die Abzugsfähigkeit der Gemeindesteuer bei der Einkommensteuer beachtet werden musste. Die effektive Gesamtsteuerbelastung ergab sich somit als:

²² Der effektive Steuersatz wurde bewusst in dieser Höhe gewählt, um die Teilnehmer implizit für die Beachtung der Steuerzahlung zu sensibilisieren.

$20\% \cdot (1 + 20\%) + 20\% + 20\% - 20\% \cdot 20\% \cdot (1 + 20\%) = 59,2\% \approx 60\%$ ²³. Die Komplexität des Steuertarifsystems nimmt folglich – bei praktisch gleicher Gesamtbelastung – von Variante A nach Variante D zu, wobei die Komplexität hier durch die Anzahl der zu berücksichtigenden Steuersätze sowie deren Interaktionen modelliert wurde. Dabei wurden die notwendigen Rechenschritte die zur Berechnung der effektiven Gesamtsteuerbelastung notwendig sind immer weiter ausgebaut und um Rechenschritte, wie sie auch im deutschen Steuersystem bei der Interaktion verschiedener Steuersätze zu finden sind, ergänzt. Dieses Vorgehen erlaubt folglich eine stufenweise Komplexitätsmodellierung. Bei den Steuersystemkarten selbst wurde versucht, die sprachlichen Formulierungen so einfach wie möglich zu halten und steuerrechtliche Fachtermini soweit möglich zu vermeiden.²⁴ Es wurde bewusst auf die Angabe komplexer Steuertariffunktionen, bspw. unter Beachtung von Progressionswirkungen, verzichtet, da dies für die Beantwortung der Forschungsfrage, aufgrund einer unverhältnismäßigen Erhöhung der Komplexität, kontraproduktiv wäre.

Nachdem die Teilnehmer das für sie geltende Steuertarifsystem erhalten hatten, wurden jedem Teilnehmer Lösungsbögen der Klausur in ausreichender Stückzahl übergeben und das Experiment wurde gemeinsam begonnen.²⁵ Die Startzeit wurde vom Experimentleiter zur späteren Auswertung der jeweiligen Arbeitszeit notiert. Jeder Teilnehmer wurde nach seiner individuellen Beendigung der Arbeitstätigkeit in einen separaten Raum geführt, wo dieser entsprechend seiner Arbeitsleistung ausbezahlt wurde und einen zusätzlichen Fragebogen bekam. Der Fragebogen beinhaltete sowohl demografische Fragen wie auch Fragen zur Zufriedenheit mit der Arbeitstätigkeit und dem entsprechenden Lohn und Steuertarifsystem. Alle Teilnehmer stimmten der Beantwortung des Fragebogens zu.²⁶ Die individuelle Arbeitszeit sowie die Arbeitsleistung (gemessen anhand der (fehlerfrei) digitalisierten Bögen) wurde exakt protokolliert.

²³ Der Rundungsfehler bei dieser Variante beträgt 0,8 Prozentpunkte und kann damit vernachlässigt werden.

²⁴ Auch die Benennung der verschiedenen Steuerarten (Einkommensteuer, Zusatzsteuer, Zuschlagsteuer, Landessteuer, Gemeindesteuer) wurde so allgemein wie möglich gehalten, um einen potenziellen Labeling- oder Affektationseffekt einer Steuer zu vermeiden. Vgl. dazu *Hundsdoerfer et al. (2010)*.

²⁵ Die Lösungsbögen waren so konzipiert, dass keine Systematik erkennbar war, die die Arbeitsaufgabe vereinfacht hätte. Sonstige Hilfsmittel waren nicht zugelassen.

²⁶ Der Fragebogen befindet sich im Anhang.

4. Ergebnisse und Diskussion

Als Maßstab für die Messung der erbrachten Arbeitsleistung kann die Arbeitszeit oder die Anzahl der digitalisierten Bögen herangezogen werden. Da die zu bewältigende Arbeitsaufgabe selbst zwar keine besonderen Fähigkeiten voraussetzte, aber dennoch ein hohes Maß an Konzentration erforderte, ist weiterhin die Produktivität und die Fehlerhäufigkeit von Interesse. Da die geleistete Arbeit vor Auszahlung des entsprechenden Arbeitsentgeltes auf Fehlerfreiheit geprüft wurde, wurden dementsprechend auch nur die fehlerfrei digitalisierten Bögen vergütet. Die folgende Tabelle stellt die gesamte erbrachte Arbeitsleistung (Anzahl der digitalisierten Bögen) sowie die entsprechende Arbeitszeit während des Experiments dar:

	Mittelwert	Median	Minimum	Maximum	Standardabweichung	Summe
Arbeitsminuten	102,84	102,50	11,00	181,00	47,27	9.873,00
Digitalisierte Bögen	74,97	66,50	4,00	199,00	48,09	7.197,00
Fehlerhafte Bögen	11,47	10,50	0,00	45,00	7,52	1.101,00
Bruttolohn	19,05	17,10	0,30	52,80	13,09	1.828,80
Nettolohn	7,62	6,84	0,12	21,12	5,24	731,52

Tabelle 2 - Gesamte Arbeitsleistung im Experiment

Die durchschnittliche Produktivität über alle Teilnehmer beträgt 0,686 Bögen pro Minute, wobei diese auf einer Intervallbreite von 0,221 Bögen pro Minute bis 1,725 Bögen pro Minute liegt und somit stark zwischen den Teilnehmern variiert. Es kann davon ausgegangen werden, dass die Produktivität bei dieser Aufgabe insbesondere von motorischen Fähigkeiten, bspw. der Hand-Augen-Koordination, abhängt. Weiterhin ist es für die Auswertung der Ergebnisse von Interesse, wie sich die Produktivität im Zeitverlauf verhält. Wenn sich die Produktivität im Zeitablauf stark ändert, würde das bedeuten, dass sich die beiden Messgrößen für das Arbeitsangebot (Arbeitszeit und Arbeitsleistung) zwischen Teilnehmern mit hohem Arbeitsangebot und Teilnehmern mit niedrigem Arbeitsangebot nicht linear verhalten. So würde ein Teilnehmer im Vergleich zu einem anderen bei doppelter Arbeitszeit nicht genau das doppelte an Arbeitsleistung aufweisen, selbst wenn die beiden Personen vom Grundsatz her mit der gleichen Produktivität starten würden. Dies könnte die Güte der Arbeitsangebotsmessung beeinflussen und muss somit berücksichtigt werden.

Um einen solchen Effekt zu überprüfen, wurden die Probanden entsprechend der gearbeiteten Zeit in sechs Gruppen eingeteilt. Um in jedem Zeitintervall ausreichend Teilnehmer für diese Darstellung zu haben, wurden die Zeitintervalle jeweils in 30-Minuten-Schritte eingeteilt. So kann eine Aussage darüber getroffen werden, inwieweit die erbrachte Arbeitsleistung auch von nicht steuerinduzierten bzw. nicht monetären Effekten abhängt. Die folgende Tabelle und Abbildung stellt diese Ergebnisse dar. Dabei bezieht sich die Bruttonproduktivität auf alle digitalisierten Bögen, die Nettoproduktivität hingegen nur auf die fehlerfreien Bögen.

Zeitintervall (in Minuten)	N	Bruttonproduktivität (Nettoproduktivität) in Bögen pro Minute				
		Mittelwert	Median	Minimum	Maximum	Standardabweichung
00 - 29	6	0,454 (0,261)	0,405 (0,261)	0,304 (0,044)	0,821 (0,500)	0,188 (0,162)
30 - 59	14	0,543 (0,375)	0,474 (0,355)	0,370 (0,140)	0,816 (0,667)	0,142 (0,146)
60 - 89	22	0,649 (0,537)	0,689 (0,562)	0,221 (0,221)	0,943 (0,725)	0,198 (0,141)
90 - 119	14	0,738 (0,641)	0,645 (0,546)	0,472 (0,366)	1,725 (1,560)	0,333 (0,320)
120 - 149	20	0,806 (0,684)	0,772 (0,653)	0,333 (0,225)	1,363 (1,181)	0,245 (0,224)
150 +	20	0,741 (0,648)	0,719 (0,633)	0,408 (0,284)	1,167 (1,024)	0,207 (0,200)
Gesamt	96	0,686 (0,565)	0,682 (0,557)	0,221 (0,044)	1,725 (1,560)	0,245 (0,240)

Tabelle 3 - Produktivitäten im Zeitablauf

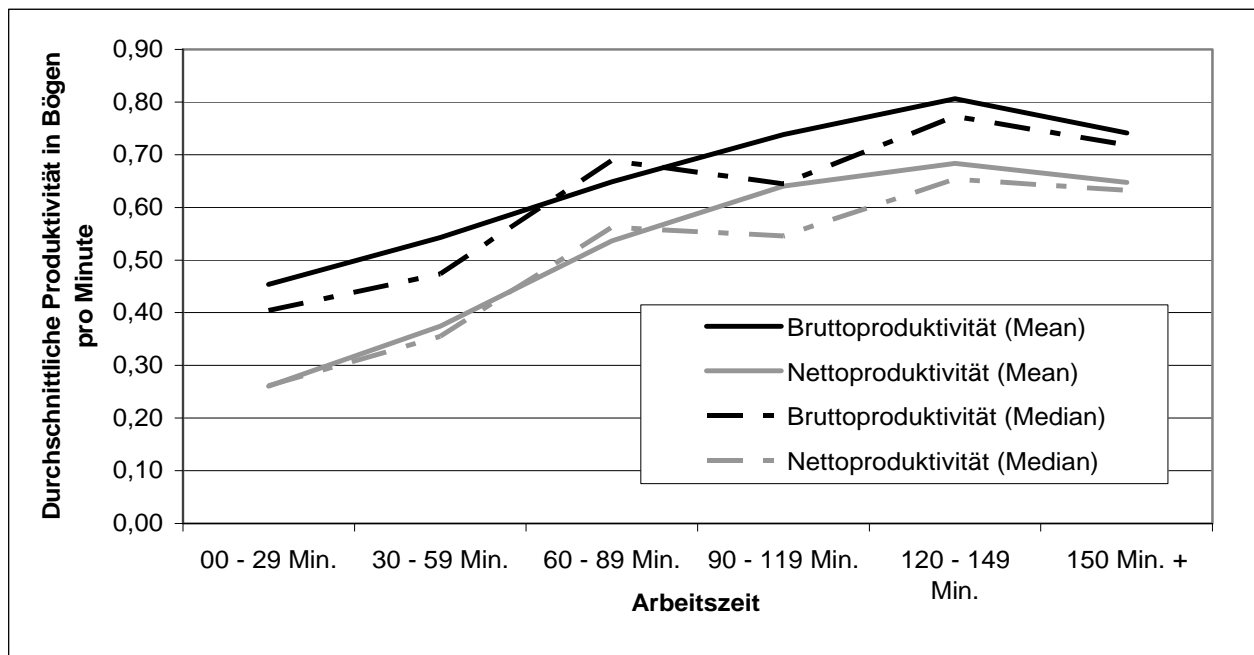


Abbildung 1 - Produktivitäten im Zeitablauf

Es wird deutlich, dass die Teilnehmer, die eine lange Zeit gearbeitet haben, im Mittel eine höhere Produktivität aufweisen als die Teilnehmer, die das Experiment schon nach kurzer Zeit beendet haben. Dieses Ergebnis könnte auf einen positiven Lerneffekt bei der Bearbeitung der Aufgabe hinweisen.²⁷ Bei den Teilnehmern mit der längsten Arbeitsdauer nimmt die durchschnittliche Produktivität dann wieder leicht ab. Ursächlich dafür können vor allem Ermüdungserscheinungen sein, da die Aufgabe ein hohes Maß an Konzentration erforderte und die Teilnehmer so nach einem anfänglichen Produktivitätsanstieg wieder einen Abfall der Produktivität aufweisen. Diese Vermutung deckt sich auch mit den Angaben der Teilnehmer im anschließend ausgefüllten Fragebogen, in dem zahlreiche Teilnehmer über Konzentrationsschwierigkeiten, Kopfschmerzen oder auftretende Verspannungen berichteten, welche nach einer langen Arbeitszeit mit zur Beendigung der Teilnahme beigetragen haben. Dieser Produktivitätseffekt wird im weiteren Verlauf der Untersuchung somit zu berücksichtigen sein.

Das zentrale Ergebnis zur Beantwortung der hier untersuchten Fragestellung ist, dass die Teilnehmer an dem Experiment länger gearbeitet und mehr Bögen digitalisiert haben, je einfacher das Steuertarifsystem hinsichtlich des Zusammenspiels verschiedener Steuersätze war. Im Mittel wurde

²⁷ Ob es sich wirklich um einen Lerneffekt handelt kann nur vermutet werden, da die Daten keinen Aufschluss darüber geben, wie sich die Produktivität der einzelnen Teilnehmer im Zeitablauf entwickelt hat. Denkbar wäre hier auch, dass die Teilnehmer, die von Anfang an eine höhere Produktivität aufweisen, ebenfalls eine höhere Arbeitsbereitschaft haben.

bei dem einfachsten Steuertarifsystem (Steuersystem A) die meiste Zeit mit der Arbeitsaufgabe verbracht und die meisten Bögen digitalisiert. Bei dem komplexen Steuertarifsystem D fällt hingegen auf, dass im Mittel weniger gearbeitet wurde.

	Mittelwert	Median	Standardabweichung	Minimum	Maximum
Arbeitsminuten	125,67	138,00	39,62	35,00	181,00
Digitalisierte Bögen	92,79	80,50	41,83	16,00	196,00
A Fehlerhafte Bögen	13,63	13,50	5,78	2,00	24,00
Bruttolohn	23,75	21,45	11,62	4,20	51,60
Nettolohn	9,50	8,58	4,65	1,68	20,64
Arbeitsminuten	117,54	113,50	38,23	54,00	171,00
Digitalisierte Bögen	80,08	74,00	45,05	20,00	198,00
B Fehlerhafte Bögen	11,54	10,50	6,41	2,00	26,00
Bruttolohn	20,56	18,45	12,51	2,40	52,80
Nettolohn	8,23	7,38	5,00	0,96	21,12
Arbeitsminuten	97,67	92,50	50,42	15,00	174,00
Digitalisierte Bögen	74,88	61,00	52,41	6,00	199,00
C Fehlerhafte Bögen	11,33	8,00	10,12	2,00	45,00
Bruttolohn	19,06	15,30	13,93	0,30	50,70
Nettolohn	7,63	6,12	5,57	0,12	20,28
Arbeitsminuten	70,50	66,00	42,05	11,00	139,00
Digitalisierte Bögen	52,13	28,50	46,09	4,00	188,00
D Fehlerhafte Bögen	9,38	8,00	6,86	0,00	23,00
Bruttolohn	12,83	6,90	12,50	0,90	51,00
Nettolohn	5,13	2,76	5,00	0,36	20,40

Tabelle 4 - Arbeitsleistung nach Gruppen

Die folgenden Abbildungen sollen die Verteilung der wesentlichen Messgrößen des Experiments in Abhängigkeit von dem jeweiligen Steuertarifsystem nochmals verdeutlichen:

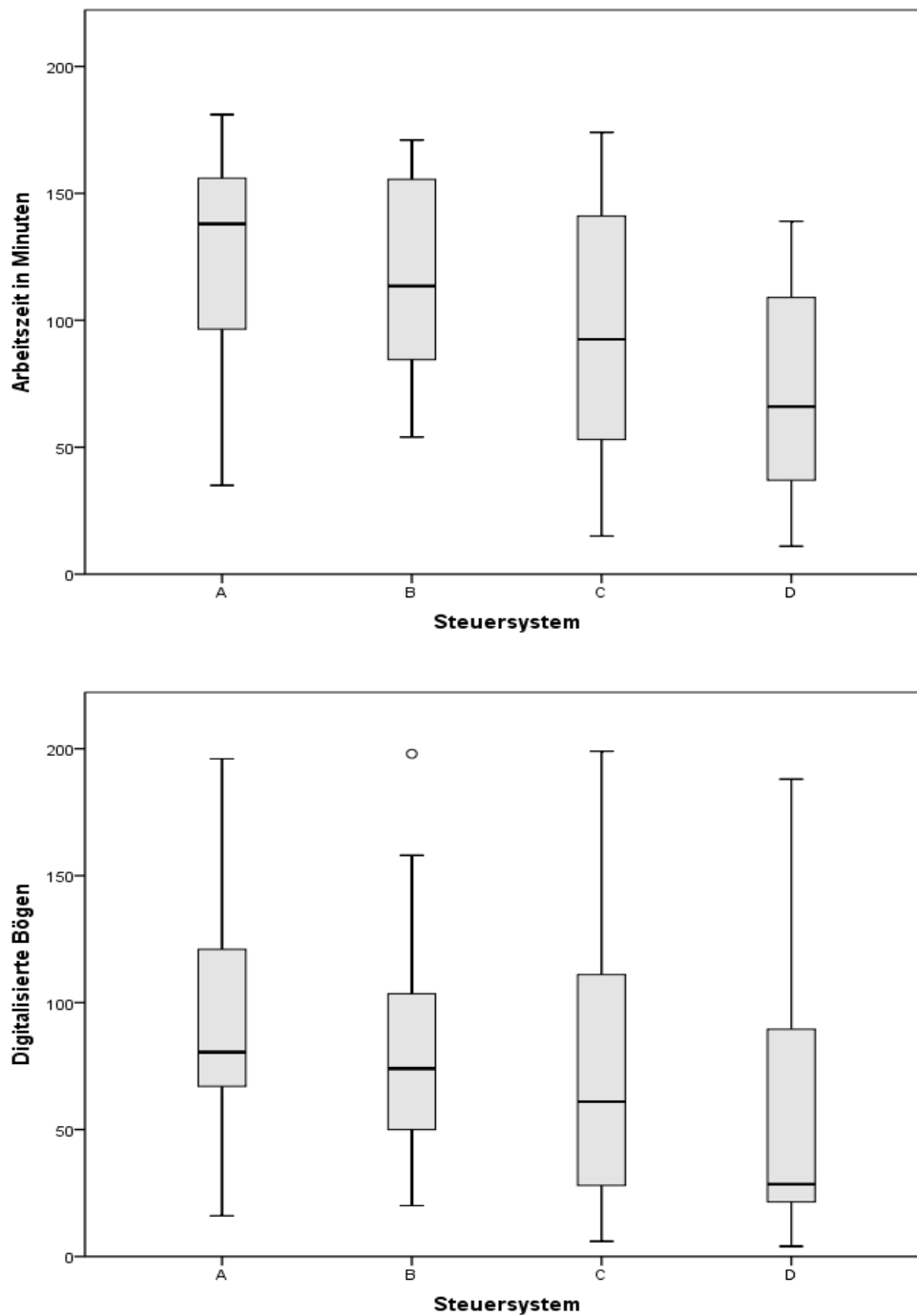


Abbildung 2 - Arbeitszeit und Arbeitsleistung

Um zu überprüfen, inwieweit die Unterschiede zwischen der Arbeitsbereitschaft unter den verschiedenen Steuertarifsystemen statistisch signifikant sind, wurden paarweise Mann-Whitney-U-

Tests²⁸ durchgeführt. Die folgende Tabelle gibt einen Überblick über die Ergebnisse. Dabei wurde sowohl auf einen Unterschied hinsichtlich der Arbeitszeit wie auch der Arbeitsleistung (Anzahl der digitalisierten Bögen) und der Produktivität getestet.²⁹ Die Tabelle enthält jeweils die p-Werte der Testverfahren.

Verglichene Gruppen	Arbeitszeit	Arbeitsleistung	Produktivität
A - B	0,470	0,265	0,180
A - C	0,051*	0,101	0,529
A - D	0,000***	0,001***	0,248
B - C	0,146	0,585	0,421
B - D	0,000***	0,015**	0,942
C - D	0,058**	0,083*	0,483

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 5 - Paarweise Gruppenvergleiche

Es wird deutlich, dass sowohl die Arbeitszeit als auch die Arbeitsleistung zwischen einigen Gruppen signifikant verschieden ist.³⁰ Während zwischen dem einfachen Steuertarifsystem (Steuersystem A) und dem um eine additive Operation erweiterten Steuertarifsystem (Steuersystem B) kein signifikanter Unterschied festzustellen ist, wird insbesondere die signifikant geringere Arbeitszeit und Arbeitsleistung bei den Teilnehmern unter Steuertarifsystem D deutlich. Alle anderen Steuertarifsystemgruppen unterscheiden sich signifikant von Steuertarifsystem D, welches den höchsten Komplexitätsgrad aufwies. Zwischen den Steuertarifsystemen B und C lässt sich des Weiteren kein signifikanter Unterschied feststellen, wobei sich diese Steuertarifsysteme durch die Einführung einer weiteren Zuschlagsteuer unterscheiden. Da sich das Arbeitsangebot zwischen Gruppe A und Gruppe B gar nicht und zwischen Gruppe A und Gruppe C nur schwach signifikant unterscheidet,

²⁸ Bei dem Mann-Whitney-U-Test handelt es sich um ein parameterfreies statistisches Testverfahren. Dieses Testverfahren überprüft, ob zwei unabhängige Verteilungen zu derselben Grundgesamtheit gehören. Aufgrund der Beschaffenheit der Daten ist die Anwendung eines parameterfreien Testverfahrens geboten. Vgl. dazu auch *Mann* und *Whitney* (1947) sowie *Wilcoxon* (1945).

²⁹ Die Produktivität dient dabei als relativer Maßstab. Damit kann untersucht werden, ob sich Gruppenunterschiede hinsichtlich Arbeitszeit und Arbeitsleistung linear zwischen den Gruppen entwickeln.

³⁰ Es wurde auch die Fehlerquote auf einen Steuertarifsystembedingten Unterschied überprüft. Hier konnten aber keine signifikanten Abweichungen identifiziert werden.

kann ein möglicher Ankereffekt des höchsten und zuerst genannten Steuersatzes (Einkommensatz) anhand dieser Ergebnisse nicht bestätigt werden.³¹

Hinsichtlich der Produktivität konnten keine Unterschiede zwischen den Steuertarifsystemgruppen gefunden werden. Es deutet nichts darauf hin, dass das Steuertarifsystem auch einen Einfluss auf die Schnelligkeit bei der Arbeitsaufgabe hat. Dies scheint der Tatsache geschuldet, dass die Produktivität vor allem von physiologischen Ursachen restringiert wird und diese von den Teilnehmern nur bedingt beeinflusst werden können. Lediglich die Arbeitszeit scheint einen messbaren Einfluss auf die Produktivitätsrate der Teilnehmer auszuüben. Die folgende Tabelle stellt die Produktivitäten nochmals getrennt nach Zeit und Steuertarifsystem dar.

Zeitintervall ³²	A	B	C	D
0 - 59 Min.	0,457	0,405	0,483	0,560
60 - 119 Min.	0,722	0,611	0,684	0,731
120 + Min.	0,755	0,728	0,857	0,788

Tabelle 6 - Produktivitäten nach Zeitintervall und Steuertarifsystem

Aus der Tabelle wird ersichtlich, dass es keine auffallenden Produktivitätsunterschiede zwischen den einzelnen Gruppen gibt. Allerdings sollte berücksichtigt werden, dass die mittlere Produktivität bei Teilnehmern mit längerer Arbeitszeit höher ist als bei Teilnehmern mit kurzer Arbeitszeit. Um diesem Umstand innerhalb der Analyse Rechnung zu tragen, wurden die oben durchgeführten Gruppenvergleiche nochmals mit einer um zeitliche Produktivitätseffekte bereinigten Arbeitszeit bzw. Arbeitsleistung durchgeführt.³³ Dazu wurde die mittlere Produktivität über alle Teilnehmer berechnet und dann die Arbeitszeit bzw. -leistung entsprechend dieser Durchschnittsproduktivität korrigiert. Dabei wurde folglich zum einen für jeden Teilnehmer die Arbeitszeit berechnet, die er bei durchschnittlicher Produktivität gebraucht hätte, um die gleiche Anzahl an Bögen zu digitalisieren, die er tatsächlich geschafft hat. Zum anderen wurde die Anzahl der Bögen ent-

³¹ Folgt man der in der überwiegenden Literatur vertretenen Meinung, dass eine Erhöhung des Nettolohns zu einer Ausweitung des Arbeitsangebots führt, so müsste ein solcher Ankereffekt ebenfalls zu einem höheren Arbeitsangebot bei Gruppe D im Vergleich zu allen anderen Gruppen führen. Auch danach kann ein Ankereffekt somit nicht bestätigt werden.

³² Aufgrund der Teilnehmerzahl wurden die Zeitintervalle bei dieser Gruppenbildung größer gefasst.

³³ Dieses Vorgehen stellt eine Normierung der mutmaßlich durch exogene Einflüsse getriebenen Produktivitätsrate dar und dient der Überprüfung der Robustheit der gefundenen Ergebnisse.

sprechend der tatsächlichen Arbeitszeit bei durchschnittlicher Produktivität angepasst. Dieses Vorgehen soll mögliche Messverzerrungen durch Produktivitätseffekte ausschließen.

Die korrigierte Arbeitszeit für Teilnehmer i ergibt sich dann als:

$$\text{Arbeitszeit}_i = \frac{\text{Bögen}_i}{\varnothing \text{ Produktivität}}$$

und die korrigierte Anzahl an Arbeitsleistung als:

$$\text{Bögen}_i = \text{Arbeitszeit}_i \cdot \varnothing \text{ Produktivität} .$$

Die folgende Tabelle stellt die mittlere Arbeitszeit sowie mittlere Arbeitsleistung in der korrigierten Fassung dar.

	Mittelwert	Median	Standardabweichung	Minimum	Maximum	
A	Arbeitsminuten	135,20	117,29	70,95	23,31	285,58
	Digitalisierte Bögen	86,25	94,71	27,19	24,02	124,22
B	Arbeitsminuten	116,68	107,82	65,64	29,14	288,49
	Digitalisierte Bögen	80,67	77,90	26,24	37,06	117,36
C	Arbeitsminuten	109,10	88,88	76,37	8,74	289,95
	Digitalisierte Bögen	67,03	63,49	34,60	10,29	119,42
D	Arbeitsminuten	75,95	41,53	41,53	5,83	273,92
	Digitalisierte Bögen	48,39	45,40	45,30	7,55	95,40

Tabelle 7 - Arbeitszeit und Arbeitsleistung in korrigierter Form

Diese Werte stellen die notwendige Arbeitszeit bzw. die vollbrachte Arbeitsleistung unter der Bedingung dar, dass alle Teilnehmer mit der gleichen Produktivität gearbeitet hätten und es keine exogenen sowie endogenen Einflüsse auf die Produktivität einzelner Teilnehmer gäbe. In diesem Fall wird unterstellt, dass die Teilnehmer ihr Arbeitsangebot entweder ausschließlich an einer festen Zeit ausgerichtet haben oder aber sich eine bestimmte Stückzahl bei der Arbeitsleistung als Ziel gesetzt haben. Effekte durch das schnellere oder langsamere Arbeiten werden dabei ausgeblendet. Es sei angemerkt, dass damit auch Produktivitätsunterschiede durch steuerinduzierte Anreize ausgeblendet werden.

Während des Experiments war weder die Kommunikation mit anderen Teilnehmern noch eine alternative Tätigkeit erlaubt, so dass es unwahrscheinlich ist, dass die Produktivität bewusst beeinflusst wurde. Es kann vermutet werden, dass die Teilnehmer während ihrer Teilnahme auch ihre volle Konzentration auf die Arbeitsaufgabe verwendet haben oder aber das Experiment abgebrochen haben. Da die Produktivitätsunterschiede somit eher durch nicht beeinflussbare Restriktionen bestimmt waren, scheint dieses Vorgehen zur Kontrolle der obigen Ergebnisse angemessen.

Hier wurden auch wieder paarweise Gruppenvergleiche mit Hilfe des Mann-Whitney-U-Tests durchgeführt. Die Ergebnisse unterscheiden sich nur schwach im Vergleich zu den vorangegangenen Tests mit der unkorrigierten Messgröße, was wiederum die Ergebnisse weiter bestätigt.

Verglichene Gruppen:	Arbeitszeit	Arbeitsleistung
A - B	0,265	0,470
A - C	0,101	0,051**
A - D	0,001***	0,000***
B - C	0,585	0,146
B - D	0,015**	0,000***
C - D	0,083*	0,058*

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 8 - Paarweise Gruppenvergleiche in korrigierter Form

Aus diesen Ergebnissen lässt sich ableiten, dass eine Erhöhung der Komplexität des Steuertarifsystems zur Folge hat, dass die Steuerpflichtigen ihr Arbeitsangebot reduzieren. Obwohl der effektive Steuersatz sowie der Brutto- und Nettolohn in allen vier Steuertarifsystemvarianten gleich ist, zeigt sich ein signifikanter Unterschied. Das Steuertarifsystem D bestand aus vier verschiedenen Steuersätzen, welche durch additive und multiplikative Verknüpfungen sowie die Abzugsfähigkeit einer Steuer bei der Berechnung einer anderen gekennzeichnet war. Dieses Geflecht aus verschiedenen Steuersätzen konnte nur mit großem Aufwand zu einem effektiven Gesamtsteuersatz zusammengefügt werden. Wenn die Teilnehmer dazu nicht in der Lage waren, bzw. ihnen diese Aufgabe als zu mühselig erschien, standen diese vor der Wahl, die Steuern entweder ganz zu vernachlässigen oder eine Vereinfachungsregel anzuwenden bzw. ihr Arbeitsangebot an einer Heuristik auszurichten. Da das Arbeitsangebot in dieser Steuertarifsystemgruppe signifikant unter

dem Angebot in den anderen Gruppen liegt, kann gefolgert werden, dass die Teilnehmer aufgrund der Unsicherheit über ihre tatsächliche Steuerzahlung und somit über ihren Nettolohn eher risikoavers reagieren und den Steuersatz als zu hoch und die Arbeit damit als weniger lohnend ansehen. Den bisherigen Untersuchungen in der Literatur über den Zusammenhang zwischen Arbeitslohn und Arbeitsangebot folgend, führt diese individuelle Einschätzung über die Steuerbelastung dann zu der hier festgestellten Reduktion des Arbeitsangebots.³⁴

Alternativ kann das eingeschränkte Arbeitsangebot auch auf eine allgemeine Unzufriedenheit zurückzuführen sein, die bei den Teilnehmern durch die Konfrontation mit vielen Steuersätzen ausgelöst wird. Dadurch könnte die Arbeitsmotivation gesenkt werden, ohne überhaupt einen Steuersatz zu berechnen bzw. zu schätzen. Erste Implikationen für den Gesetzgeber sind hier bereits ersichtlich. Offensichtlich bewirkt nicht nur die Höhe der effektiven Steuerbelastung eine Anpassungsreaktion bei den Zensiten, sondern auch der Grad der Komplexität. Komplexe Steuertarifsysteme wirken dabei klar arbeitsangebotsreduzierend. Die Strategie, einzelne kleine Steuerbelastungen in bzw. hinter einem komplexen Steuersystem zu verstecken, um so das Belastungsgefühl der Zensiten zu senken, scheint vor diesem Hintergrund keine Erfolg versprechende Strategie zu sein, da die Komplexität selbst, zumindest in einem gewissen Rahmen, zu Steuerwiderstand führt. Dieser äußert sich in diesem Experiment durch eine Reduzierung des Arbeitsangebots.

Zusätzlich zu den Ergebnissen aus dem Experiment wurden die Teilnehmer im Anschluss noch gebeten, einen Fragebogen zu ihren Erfahrungen während des Experiments auszufüllen. Insbesondere wurden die Teilnehmer gefragt, ob sie die Vergütung fair fanden bzw. wie sie diese beurteilen, welchen Einfluss die Besteuerung und der entsprechende Nettolohn auf ihre Arbeitsbereitschaft hatten und warum sie das Experiment beendet haben. Außerdem wurden die Teilnehmer gebeten, den für sie geltenden Steuersatz noch einmal zu nennen, sofern sie diesen direkt ablesen konnten (Steuersystem A) oder diesen vor dem Experiment bereits berechnet hatten. Wenn sie den Steuersatz noch nicht vor Beginn des Experiments berechnet hatten, sollten sie dies jetzt nachholen. Die folgende Tabelle fasst die Mittelwerte der relevanten Fragen zusammen. Fragen hinsichtlich der Einschätzung der Tätigkeit, des Einflusses von Steuern und Nettolohn, der Fairness sowie der Steuerrechtskenntnisse wurde jeweils auf einer Skala von eins bis zehn abgefragt. Die entsprechenden Fragen finden sich im Anhang.

³⁴ Vgl. Fußnote 14.

Steuersystem:	A	B	C	D
Wie angenehm fanden Sie die Tätigkeit? (1 = sehr unangenehm; 10 = sehr angenehm)	4,96	4,21	4,54	3,75
Anteil der Teilnehmer, die ihren Steuersatz vor Beginn der Tätigkeit berechnet haben.	88%	83%	58%	38%
Nennen Sie uns jetzt bitte nochmals den Steuersatz bzw. berechnen sie diesen jetzt. ³⁵	58,33% (60%) (0,0816)	57,35% (60%) (0,0520)	59,06% (60%) (0,0804)	63,92% (64%) (0,1103)
Welchen Einfluss hatte die Steuer auf ihr Arbeitsangebot? (1 = gar keinen Einfluss; 10 = sehr großen Einfluss)	6,25	5,63	5,83	6,79
Nennen Sie uns jetzt bitte nochmals den Nettolohn bzw. berechnen Sie diesen jetzt.	12,78	14,41	13,71	13,32
Welchen Einfluss hatte der Nettolohn auf ihr Arbeitsangebot? (1 = gar keinen Einfluss; 10 = sehr großen Einfluss)	7,25	6,54	6,04	5,88
Wie fair fanden Sie die Vergütung? (1 = gar nicht fair; 10 = sehr fair)	4,58	4,00	3,92	4,54
Wie hoch schätzen Sie ihre eigenen Steuerrechtskenntnisse ein? (1 = keine Kenntnisse; 10 überdurchschnittliche Kenntnisse)	4,50	4,04	4,17	3,79

Tabelle 9 - Fragen zum Experiment

Die Tabelle zeigt, dass in Gruppe A der Steuersatz von 88% der Teilnehmer vor Beginn des Experiments berechnet bzw. beachtet wurde, wohingegen dieser Wert mit steigender Komplexität abnimmt.³⁶ In Gruppe D haben lediglich 38% der Teilnehmer vor dem Experiment ihren Steuersatz berechnet. Da bei Gruppe A der effektive Steuersatz explizit auf der Steuersystemkarte angegeben war, wäre hier ein Ergebnis von 100% zu erwarten. Diese Differenz ist wohl darauf zurückzu-

³⁵ Die Werte in Klammern geben die Mediane sowie die Standardabweichung bei dieser Frage an.

³⁶ Diese Frage zielte im Fragebogen darauf ab, ob der Steuersatz vor Aufnahme der Arbeitstätigkeit berechnet wurde. Da bei Gruppe A der effektive Steuersatz direkt auf der Steuersystemkarte angegeben wurde, sollte diese Frage bei der Gruppe Aufschluss darüber geben, ob dieser Steuersatz überhaupt von den Teilnehmern beachtet bzw. wahrgenommen wurde.

führen, dass 12% der Probanden die Frage verneint haben, da diese dem Steuersatz vor dem Experiment offensichtlich einfach keine Beachtung geschenkt haben. Auch bei der Berechnung des Steuersatzes liegt der Mittelwert in Gruppe A nicht genau bei 60%. Dies ist einem Probanden geschuldet, der trotz der konkreten Angabe des effektiven Steuersatzes auf der Steuersystemkarte einen Wert von 20% als effektive Steuerbelastung im Fragebogen angegeben hat. Deshalb wurden in die Tabelle ebenfalls die Mediane bei dieser Frage angegeben. Insbesondere in Gruppe D fällt auf, dass der Steuersatz im Durchschnitt überschätzt wurde. Hier liegen insgesamt 12 der 24 Gruppenteilnehmer mit ihrer Schätzung über den eigentlichen 60% Steuersatz. Drei Probanden haben hier einen Wert von 80% angegeben, was der Summe aller auf der Karte angegebenen Steuersätze entspricht und als Anzeichen für eine Vereinfachungsheuristik bei der Kalkulation des Steuersatzes gewertet werden kann. Auch die angegebene Standardabweichung zeigt, dass bei Gruppe D die angestellten Berechnungen bzw. Schätzungen weiter um den Mittelwert streuen, was ebenfalls als ein Zeichen für die Berechnungsschwierigkeiten gedeutet werden kann. Weiterhin fällt auf, dass nur in Gruppe D die Wichtigkeit des Nettolohns bei der Arbeitsangebotsentscheidung niedriger als die Wichtigkeit der Steuer selbst ist. Ursächlich dafür könnte die Tatsache sein, dass die Probanden hier den Nettolohn nicht oder nur in Ansätzen berechnen konnten, was wiederum dazu führt, dass alleine das Steuertarifsystem deren Arbeitsangebot, bspw. aufgrund der Konfrontation mit vier verschiedenen Steuerarten, gesenkt hat. Die Tätigkeit selbst scheint von den Teilnehmern in den ersten drei Steuertarifsystemgruppen auch im Mittel als angenehmer empfunden worden zu sein, was ebenfalls ein Indiz für die abschreckende Wirkung des komplexen Steuertarifsystems ist. Da die Probanden zufällig auf die Steuertarifsysteme verteilt wurden, scheint hier ebenfalls das Steuertarifsystem verantwortlich für die Zufriedenheit mit der konkreten Form der Arbeitsaufgabe zu sein.

Zur Bestätigung dieser Ergebnisse wurden wiederum parameterfreie Testverfahren durchgeführt. Damit konnten signifikante Unterschiede auf einem Signifikanzniveau von 5% bei der Frage nach der Höhe des Steuersatzes, bei der Frage, ob der Steuersatz oder der Nettolohn vorher berechnet wurden und bei der Frage, wie angenehm die Tätigkeit empfunden wurde, festgestellt werden. Bei den anderen Fragen gibt es keine signifikanten Unterschiede zwischen den einzelnen Steuertarifsystemgruppen.

Zum Schluss der Untersuchung sollen die Einflüsse des Steuertarifsystems sowie weiterer Größen auf das Arbeitsangebot mittels einer OLS Regression genauer quantifiziert werden. Eine Regression eignet sich dafür besonders, da so der Einfluss mehrerer erklärender Variablen gleichzeitig identi-

fiziert und analysiert werden kann. Als abhängige Variable wird dabei zum einen die Arbeitszeit und zum anderen die Arbeitsleistung (digitalisierte Bögen) verwendet. Die zu schätzende Regressionsgleichung ergibt sich somit als:

$$y_i = \beta_0 + \underbrace{\beta_1 \text{Dummy}_i^B + \beta_2 \text{Dummy}_i^C + \beta_3 \text{Dummy}_i^D}_{\text{Dummyvariablen für die Steuersystemgruppen}} + \underbrace{\beta_k \mathbf{X}_{i,k}}_{\text{Kontrollvariablen}} + \varepsilon_i,$$

wobei die Matrix \mathbf{X} die Kontrollvariablen: „Einfluss der Besteuerung“, „Zufriedenheit mit der Tätigkeit“, „Fairness der Vergütung“, „Steuerrechtskenntnisse“, „Alter“, „Geschlecht“, „Einkommensklassen“ sowie „Studium der Betriebswirtschaftslehre“ enthält.

Die folgenden Tabellen stellt die Ergebnisse der Regression dar. Dabei wurde die abhängige Variable jeweils in ihrer tatsächlichen Form wie auch in der um Produktivitätseffekte korrigierten Form verwendet.

Abhängige Variable	Tatsächliche Arbeitszeit		Korrigierte Arbeitszeit	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	97,683***	34,105	91,201**	45,766
Steuersystem B	-11,279	12,576	-26,183	16,875
Steuersystem C	-29,760**	12,011	-30,748*	16,118
Steuersystem D	-48,646***	12,449	-57,774***	16,706
Einfluss der Besteuerung	-3,066*	1,733	-6,508***	2,325
Zufriedenheit mit der Tätigkeit	6,630***	2,408	11,328***	3,232
Fairness der Vergütung	1,356	2,326	4,356*	3,122
Steuerrechtskenntnisse	-0,370	2,254	-4,090	3,025
Alter	1,041	1,250	1,903	1,678
Geschlecht	-4,914	9,022	3,223	12,106
Einkommen 501 - 1.000 Euro	-16,464	10,519	-28,974**	14,115
Einkommen größer 1.000 Euro	-17,519	16,841	-29,064	22,599
Studium der Betriebswirtschaftslehre	3,357	11,765	10,948	15,788

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 10 - OLS Regression, Abhängige Variable: Arbeitszeit, Erweiterte Spezifikation

Da vermutet werden kann, dass die Kontrollvariablen „Einfluss der Besteuerung“, „Zufriedenheit mit der Tätigkeit“, „Fairness der Vergütung“ und „Steuerrechtskenntnisse“ in dem Modell endogen sind (da diese durch das Steuertarifsystem beeinflusst werden), wird diese Regression zur Überprüfung der Ergebnisse nochmals unter Ausschluss dieser Variablen durchgeführt.

Abhängige Variable	Tatsächliche Arbeitszeit		Korrigierte Arbeitszeit	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	110,502***	31,202	119,023**	47,119
Steuersystem B	-16,021	13,323	-31,997	20,120
Steuersystem C	-32,945**	12,782	-34,178*	19,303
Steuersystem D	-58,833***	12,823	-72,756***	19,365
Alter	1,394	1,247	1,792	1,883
Geschlecht	-7,076	9,371	-2,996	14,151
Einkommen 501 - 1.000 Euro	-19,620*	11,178	-32,461*	16,880
Einkommen größer 1.000 Euro	-15,431	18,043	-23,445	27,247
Studium der Betriebswirtschaftslehre	3,700	11,395	5,806	17,208

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 11 - OLS Regression, Abhängige Variable: Arbeitszeit, Basisspezifikation

Die Regression zeigt, dass die Steuertarifsysteme C und D einen signifikanten Einfluss auf das Arbeitsangebot relativ zur Referenzgröße „Steuersystem A“ aufweisen. Die Zugehörigkeit zu einem dieser Steuertarifsysteme führt somit zu einer Senkung des Arbeitsangebotes. Dieses Ergebnis deckt sich mit den oben durchgeführten Gruppenvergleichen. Außerdem wird ersichtlich, dass Teilnehmer, die im Fragebogen einen höheren Einfluss der Besteuerung bei ihrer Entscheidungsfindung angegeben haben, etwas weniger gearbeitet haben als Teilnehmer, die der Besteuerung nur einen geringen Einfluss beigemessen haben. Dieses Ergebnis war insoweit zu erwarten und bestätigt den gefundenen Einfluss der Besteuerung auf die Arbeitsangebotsentscheidung weiter. Ebenfalls übt die Zufriedenheit mit der Art der Arbeitstätigkeit einen hoch signifikanten und positiven Einfluss auf die Arbeitszeit aus. Beim Einkommen wurde die Einkommensklasse „bis 500 Euro monatlich verfügbares Einkommen“ als Referenzkategorie gewählt. Nur für die korrigierte Arbeitszeit zeigt sich hier ein signifikanter negativer Einfluss bei der Einkommensklasse 501 – 1.000 Euro. Das Ergebnis für die andere Einkommensklasse ist dagegen nicht signifikant. Dieses Ergebnis ist auch insoweit schlüssig, als dass die Studierenden, die mehr Geld zur Verfügung haben, die Verdienstmöglichkeiten in dem Experiment vermutlich nicht als so hoch einschätzen, was wiederum deren Arbeitszeitangebot im Mittel senkte. Sofern diese ihr verfügbares Einkommen selbst verdienen, ist auch

von höheren Opportunitätskosten auszugehen. Für die tatsächliche Arbeitszeit ist dieser Einfluss zwar ebenfalls negativ, aber in der erweiterten Spezifikation nicht signifikant. Die Variablen „Alter“, „Geschlecht“ sowie „Studium der Betriebswirtschaftslehre“ üben keinen signifikanten Einfluss auf die Arbeitszeit aus. Insgesamt liefern beide Regressionen sowohl für die erweiterte Spezifikation wie auch für die Basisspezifikation ein ähnliches Ergebnis. Im Folgenden werden die gleichen Regressionen für die Messung der Arbeitsleistung (Bögen pro Minute) durchgeführt:

Abhängige Variable	Tatsächlich digitalisierte Bögen		Korrigierte Bögen	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	62,596**	31,410	67,042***	23,407
Steuersystem B	-17,970	11,582	-7,741	8,631
Steuersystem C	-21,102*	11,062	-20,427**	8,244
Steuersystem D	-39,652***	11,466	-33,387***	8,544
Einfluss der Besteuerung.	-4,467***	1,596	-2,104*	1,189
Zufriedenheit mit der Tätigkeit	7,774***	2,218	4,550***	1,653
Fairness der Vergütung	2,990	2,143	,930	1,597
Steuerrechtskenntnisse	-2,807	2,076	-,254	1,547
Alter	1,306	1,151	,714	,858
Geschlecht	2,212	8,309	-3,373	6,192
Einkommen 501 - 1.000 Euro	-19,885**	9,688	-11,299	7,219
Einkommen größer 1.000 Euro	-19,947	15,510	-12,024	11,558
Studium der Betriebswirtschaftslehre	7,514	10,836	2,303	8,075

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 12 - OLS Regression, Abhängige Variable: Arbeitsleistung, Erweiterte Spezifikation

Abhängige Variable	Tatsächlich digitalisierte Bögen		Korrigierte Bögen	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	81,688**	32,338	75,841***	21,414
Steuersystem B	-21,960	13,809	-10,995	9,144
Steuersystem C	-23,457**	13,248	-22,613**	8,773
Steuersystem D	-49,934***	13,290	-40,379***	8,801
Alter	1,230	1,292	,957	,856
Geschlecht	-2,055	9,712	-4,856	6,431
Einkommen 501 - 1.000 Euro	-22,279**	11,585	-13,465*	7,671
Einkommen größer 1.000 Euro	-16,091	18,700	-10,591	12,383
Studium der Betriebswirtschaftslehre	3,984	11,810	2,538	7,821

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 13 - OLS Regression, Abhängige Variable: Arbeitsleistung, Basisspezifikation

Auch für die Arbeitsleistung lässt sich ein ähnliches Ergebnis wie für die Arbeitszeit ableiten. Hier hat die Erhöhung der Komplexität des Steuertarifsystems ebenfalls einen signifikant negativen Einfluss auf das Arbeitsangebot. Dagegen wirkt sich auch hier die Zufriedenheit mit der Arbeitsaufgabe positiv auf das Arbeitsangebot aus. Der Einfluss der Besteuerung hat ebenfalls wieder einen leicht negativen Einfluss. Insgesamt decken sich diese Ergebnisse mit den vorherigen Regressionsergebnissen. Auch die zur Überprüfung durchgeführte Basisspezifikation zeigt wieder ein nahezu identisches Bild wie die erweiterte Spezifikation.

Es wird somit deutlich, dass nicht nur der effektive Steuersatz allein einen Einfluss auf das Verhalten der Zensiten ausübt. Auch der Komplexitätsgrad des Tarifsystems scheint einen maßgeblichen und quantifizierbaren Einfluss auszuüben, was hier am Beispiel der Arbeits-Freizeit-Entscheidung gezeigt werden konnte. Die Annahme von exakt kalkulierenden Steuerpflichtigen in theoretischen Modellen scheint vor diesem Hintergrund zumindest dann mehr als fraglich, wenn aus diesen Modellen auf das tatsächliche Verhalten in der Realität geschlossen werden soll. Die Teilnehmer haben sich offensichtlich Heuristiken und Vereinfachungsregeln bedient, welche ihr Verhalten maßgeblich determiniert haben. Auch die Richtung dieser Verhaltensbeeinflussung ist dabei deutlich erkennbar. So führt ein komplexes Steuertarifsystem zu einer Reduzierung des Arbeits-

angebotes und nicht etwa zu einer vereinfachten Orientierung an Vor-Steuer-Werten, welche dann zu einer Angebotsausweitung führen müssten.

Auch für die Steuerpolitik sind diese Ergebnisse von Bedeutung. Das Experiment hat gezeigt, dass es eben nicht nur auf den tatsächlichen effektiven Steuersatz ankommt. Vielmehr kann eine vereinfachte Darstellung des Steuertarifsystems dabei helfen, die richtigen Anreize bei den Zensiten zu setzen und komplexitätsbedingte Fehlanreize zu vermeiden. Diesem Punkt sollte in der Steuerreformdiskussion mehr Beachtung geschenkt werden. Den effektiven Steuersatz zu senken, dafür aber immer komplexere Regelungen zu implementieren, scheint vor diesem Hintergrund nicht das angemessene Vorgehen zu sein.

5. Zusammenfassung und Ausblick

Ziel des vorliegenden Beitrags war es, den Einfluss steuerlicher Tarifkomplexität auf die Arbeits-Freizeit-Entscheidung von Steuerpflichtigen zu untersuchen. Es wurde zu diesem Zweck ein Experiment durchgeführt. Dabei wurden reale monetäre Anreize gesetzt, um so das Verhalten der Teilnehmer unter realitätsnahen Bedingungen zu beobachten. Die Teilnehmer des Experiments wurden mit einer Arbeitsaufgabe konfrontiert, bei welcher diese die Möglichkeit hatten, einen Bruttolohn in Abhängigkeit von ihrer Arbeitsleistung zu erzielen. Weiterhin bekamen die Teilnehmer eines von vier verschiedenen komplexen Steuertarifsystemen zugeteilt, welches für die Besteuerung des von ihnen verdienten Arbeitslohnes zur Anwendung kam. Die Komplexität wurde dabei durch das Zusammenspiel verschiedener einzelner Steuersätze modelliert, welche additiv oder multiplikativ verknüpft werden mussten, um die effektive Gesamtsteuerbelastung zu berechnen. Der effektive Steuersatz war bei allen Varianten gleich, was dazu führte, dass auch der Nettolohn gleich war. Es konnte gezeigt werden, dass die Gruppe mit dem höchsten Steuerkomplexitätsgrad signifikant weniger arbeitete als die Gruppen mit niedrigeren Steuerkomplexitätsgraden, wobei sich auch dort Abstufungen beim Arbeitsangebot zeigen ließen. Das Arbeitsangebot wurde über die selbst gewählte Arbeitsdauer und über die jeweils erbrachte Arbeitsleistung modelliert. Auch mögliche exogene Produktivitätsunterschiede wurden bei der Analyse berücksichtigt. Im Ergebnis konnte gezeigt werden, dass Steuerpflichtige auf komplexe Tarifsysteme mit einer Reduktion ihres Arbeitsangebots reagieren. Tarifkomplexität führt demnach zu einer systematischen Abweichung vom theoretischen Modell eines exakt kalkulierenden Steuerpflichtigen. Dieses Ergebnis besitzt sowohl für die theoretische Forschung als auch für die steuerpolitische Praxis Relevanz und sollte bei der Erarbeitung von theoretischen Modellen und auch bei der Steuerreformdiskussion beachtet

werden. Vor allem die These, dass der Staat Abgaben durch die Komplexität des Steuersystems verschleiern und somit den Steuerwiderstand mindern kann, kann hiermit widerlegt werden.

Im Folgenden sollen noch die Grenzen dieser Untersuchung sowie zukünftiges Forschungspotenzial aufgezeigt werden. Zur Überprüfung der Ergebnisse sollte die Untersuchung mit einer größeren Anzahl an Personen durchgeführt werden. Auch die Auswahl der Probanden, welche hier eine homogene Gruppe darstellen, könnte noch hinsichtlich ihrer Repräsentativität ausgebaut werden. Nicht modelliert wurden Lerneffekte, die in der Praxis durch die wiederholte Beschäftigung mit Ertragsteuern auftreten könnten. Allerdings befindet sich weltweit die Einkommensbesteuerung seit jeher im Fluss, und bspw. in Deutschland erhalten neue Tarifnormen nahezu jährlich Einzug in das Einkommensteuerrecht. Deshalb scheinen Lerneffekte bei der Konzeption dieser Fragestellung keine wesentliche Rolle zu spielen. Wenn sich ein Steuerpflichtiger in der Realität die Mühe macht, seine Steuerbelastung exakt zu berechnen, um auf dieser Grundlage seine ökonomischen Entscheidungen zu treffen, ist es eher unwahrscheinlich, dass diese Rechnung im nächsten Jahr noch genauso aussehen wird. Es sei dabei nur an die Einführung verschiedener Schedulisierungsnormen in der letzten Zeit, wie bspw. der Abgeltungsteuer oder der Thesaurierungsbegünstigung, gedacht.

Anhang

Anleitung

Mit der Teilnahme an diesem Experiment haben Sie die Möglichkeit, Geld zu verdienen. Die Höhe des Verdienstes am Ende des Experiments richtet sich nach ihrer persönlichen Anstrengung. Bitte lesen Sie diese Anleitung aufmerksam durch. Bei Fragen wenden Sie sich bitte an die Experimentleitung.

Das Ziel dieses Experiments ist es, Informationen über Ihr persönliches Arbeitsangebot zu erhalten. Zu diesem Zweck werden Sie mit einer realen Arbeitssituation konfrontiert, bei welcher Sie Geld verdienen können. Zur Vergleichbarkeit der Daten verschiedener Teilnehmer handelt es sich bei der Arbeitsaufgabe um eine Tätigkeit, welche keine Vorkenntnisse oder speziellen Fähigkeiten erfordert. Wir möchten Sie darauf hinweisen, dass Gespräche mit anderen Teilnehmern, das Verlassen des Platzes sowie die Nutzung des PCs für private Zwecke während der Durchführung des Experiments nicht erlaubt sind. Nachdem Sie diese Anleitung gelesen haben, erhalten Sie mehrere Lösungsbögen einer Klausur, welche zu digitalisieren sind. Dazu finden Sie auf Ihrem Rechner eine vorgefertigte Exceldatei, in welche die Lösungsbögen zu übernehmen sind. Wie das Digitalisierungsverfahren konkret durchgeführt werden soll, wird Ihnen nach dem Lesen dieser Anleitung exemplarisch vom Experimentleiter vorgeführt.

*Abhängig von der Anzahl der fehlerfrei digitalisierten Bögen erhalten Sie am Ende des Experiments ihre Gehaltszahlung. Für jeden digitalisierten Bogen erhalten Sie einen **Bruttolohn in Höhe von 30 Cent**. Wenn Sie also beispielsweise durchschnittlich einen Bogen pro Minute digitalisieren, erhalten Sie ein Gehalt von 18 Euro brutto pro Stunde; bei durchschnittlich 1,5 Bögen pro Minute erhalten Sie ein Gehalt von 27 Euro brutto pro Stunde. Bitte beachten Sie: Von Ihrem Bruttolohn müssen Sie noch Steuerzahlungen leisten, nur der Rest (Nettolohn) wird Ihnen am Ende des Experiments bar ausgezahlt. Die Höhe der zu zahlenden Steuer bzw. den entsprechenden Steuersatz entnehmen Sie bitte der Steuertarifsystemkarte, welche Ihnen noch separat ausgehändigt wird. Lesen Sie auch diese bitte vor Beginn der Tätigkeit sorgfältig durch!*

Sie entscheiden selbst über Ihr Arbeitsangebot! Das bedeutet, dass Sie so lange und soviel arbeiten können wie Sie wollen. Sie können das Experiment also jederzeit beenden und erhalten im Anschluss daran Ihre Gehaltsauszahlung in Abhängigkeit von Ihrer erbrachten Arbeitsleistung. Bitte melden Sie sich beim Experimentleiter, wenn Sie das Experiment beenden wollen.

Viel Vergnügen!

Steuersystemkarten

Bei Aushändigung der Karten an die Probanden war jede Karte auf die gleiche Größe normiert.

A

(1) Die Steuer auf das Einkommen beträgt 60%.

B

(1) Die Steuer auf das Einkommen (Einkommensteuer) beträgt 40%.

(2) Zusätzlich wird eine Zusatzsteuer auf das Einkommen in Höhe von 20% erhoben.

C

(1) Die Steuer auf das Einkommen (Einkommensteuer) beträgt 40%.

(2) Zusätzlich wird eine Zuschlagsteuer auf die Einkommensteuer (nicht auf das Einkommen selbst!) in Höhe von 20% erhoben.

(3) Außerdem fällt eine Landessteuer in Höhe von 12% auf das Einkommen an.

D

(1) Die Steuer auf das Einkommen (Einkommensteuer) beträgt 20%.

(2) Des Weiteren wird eine Zuschlagsteuer auf die Einkommensteuer (nicht auf das Einkommen selbst!) in Höhe von 20% erhoben.

(3) Zusätzlich fällt eine Landessteuer in Höhe von 20% auf das Einkommen an.

(4) Weiterhin fällt eine Gemeindesteuer von ebenfalls 20% auf das Einkommen an.

(5) Die Gemeindesteuer vermindert die Höhe des Einkommens, auf welches die Einkommensteuer erhoben wird (Abzugsfähigkeit der Gemeindesteuer bei der Berechnung der Einkommensteuer).

Fragen aus dem Fragebogen:

1. Wie fanden Sie die Tätigkeit, die Sie während des Experimentes ausüben sollten auf einer Skala von 1 = „sehr unangenehm“ bis 10 = „sehr angenehm“?

sehr unangenehm

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr angenehm

2. Haben Sie vor Beginn der Arbeitstätigkeit Ihren Steuersatz ermittelt bzw. dieses versucht?

- Ja.
 Nein.

3. Nennen Sie uns bitte den für Sie geltenden effektiven (Gesamt-)Steuersatz jetzt noch einmal, bzw. versuchen Sie diesen bitte jetzt zu berechnen, wenn Sie dies noch nicht während des Experiments getan haben.

Steuersatz (in Prozent): _____%.

4. Welchen Einfluss hatte die erhobene Steuer auf Ihre Arbeitsbereitschaft auf einer Skala von 1 = „gar keinen Einfluss“ bis 10 = „sehr großen Einfluss“.

gar keinen Einfluss

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr großen Einfluss

5. Haben Sie vor Beginn des Experiments Ihren Nettolohn pro Bogen ermittelt?

- Ja.
 Nein.

6. Wie wichtig war für Sie die Höhe des Nettolohns bei der Entscheidung zu arbeiten auf einer Skala von 1 = „ganz unwichtig“ bis 10 = „sehr wichtig“.

ganz unwichtig

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr wichtig

7. Wie fair fanden Sie die Vergütung für die von Ihnen geforderte Tätigkeit auf einer Skala von 1 = „gar nicht fair“ bis 10 = „sehr fair“.

gar nicht fair

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr fair

8. Wie schätzen Sie Ihre eigenen Steuerrechtskenntnisse auf einer Skala von 1 = „keine Kenntnisse“ bis 10 = „überdurchschnittliche Kenntnisse“ ein?

Keine Kenntnisse

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Überdurchschnittliche Kenntnisse

9. Wenn Sie gerade studieren, nennen Sie uns bitte Ihren Studiengang sowie ihr aktuelles Fachsemester:

- Studiengang:_____.
- Fachsemester:_____.

10. Haben Sie in Ihrem Studium bereits Veranstaltungen im Vertiefungsgebiet „Steuerlehre“ besucht?

- Ja.
- Nein.

11. Geschlecht:

- männlich
- weiblich

12. Alter:

_____ Jahre

13. Wie hoch ist ungefähr Ihr monatliches zur Verfügung stehendes Einkommen?

- < 500 €
- 501 € – 1.000 €
- 1.001 € – 1.500 €
- 1.501 € – 2.000 €
- > 2.000 €

Literaturverzeichnis

- Alm, J. (2010), Testing Behavioral Public Economics Theories in the Laboratory. *National Tax Journal*, 63 (4), 635-658.
- Alm, J., B. Jackson, M. McKee (1992), Institutional Uncertainty and Taxpayer Compliance. *American Economic Review*, 82 (4), 1018-1026.
- Atkinson, A. B., J. E. Stiglitz, (1976), The design of tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6 (1), 55-75.
- Bach, S. (2011), Kirchhof macht sich das Steuerrecht einfach. Kommentar, *DIW Wochenbericht*, 28/2011, 13. Juli 2011.
- Bartolome, C. A. de (1995), Which tax rate do people use: Average or marginal? *Journal of Public Economics*, 56, 79-96.
- Beck, P. J., J. S. Davis, W.-O. Jung (1991), Experimental evidence on taxpayer reporting under uncertainty. *Accounting Review*, 66 (3), 535-558.
- Bizer, K. (2004), Das Verhältnis von Komplexität und Steuerehrlichkeit: experimentelle Evidenz, in: Bizer, K., A. Falk, J. Lange (Hrsg.), *Am Staat vorbei - Partizipation, Fairness und Transparenz contra Steuerhinterziehung*. Berlin, 59-90.
- Bizer, K. (2008), *Steuervereinfachung und Steuerhinterziehung – eine experimentelle Analyse zur Begründung von Steuereinfachheit*. Duncker & Humblot, Berlin.
- Blaufus, K., J. Bob., J. Hundsdoerfer, D. Kiesewetter, J. Weimann (2009), It's All About Tax Rates – An Empirical Study of Tax Perception. *arqus-Diskussionsbeitrag*, Nr. 106.
- Blaufus, K., R. Ortlieb (2009), Is Simple Better? A Conjoint Analysis of the Effects of Tax Complexity on Employee Preferences Concerning Company Pension Plans. *Schmalenbach Business Review*, 61, 60-83.
- Blumkin, T., B. J. Ruffle, Y. Ganun (2008), Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment with Real Goods. *Working Paper*.

- Boylan, S. J., P. J. Frischmann (2006), Experimental Evidence on the Role of Tax Complexity in Investment Decision. *Journal of the American Taxation Association*, 28, 69-88.
- Break, G. F. (1957), Income Taxes and Incentives to Work: An Empirical Study. *The American Economic Review*, 47 (5), 529-549.
- Chetty, R., A. Looney, K. Kroft (2009), Salience and taxation: theory and evidence. *The American Economic Review*, 99, 1145-1177.
- Chetty, R., E. Saez (2009), Teaching the tax code: Earning responses to an Experiment with EITC Recipients. *NBER Working Paper*, Nr. 14836.
- Congdon, W. J., J. R. Kling, S. Mullainathan (2009), Behavioral Economics and Tax Policy. *National Tax Journal*, 62, 375-386.
- Epley, N., T. Gilovich (2006), The Anchoring-and-Adjustment Heuristic: Why the Adjustments Are Insufficient. *Psychological Science*, 17 (4), 311-318.
- FDP (2010), Für faire Finanzbeziehungen zwischen Bürger und Staat. Ob als Steuerzahler oder als Transferempfänger – Leistung muss sich immer lohnen. Beschluss des 61. Ordentlichen Bundesparteitags der FDP.
- Fochmann, M., J. Weimann, K. Blaufus, J. Hundsdoerfer, D. Kiesewetter (2010), Gross-wage illusion in a real effort experiment. *FEMM Working Paper*, No. 09.
- Fujii, E. T., C. B. Hawley (1988), On the Accuracy of Tax Perceptions. *Review of Economics & Statistics*, 70, 344-347.
- Gamage, D., A Hayashi, B. K. Nakamura (2010), Experimental Evidence of Tax Framing Effects on the Work/Leisure Decision. *Working Paper*.
- Gensemer, B. L., J. A. Lean, W. B. Neenan (1965), Awareness of Marginal Income Tax Rates Among High-Income Taxpayers. *National Tax Journal*, 18, 258-267.
- Gigerenzer, G., W. Gaissmaier (2011), Heuristic decision making. *Annual Review of Psychology*, 62, 451-482.

- Harberger, A. C. (1964), The Measurement of Waste. *American Economic Review*, 54 (3), 58-76.
- Hausman, J. (1981), Labor Supply, in: Aaron, H., J. Pechman (Hrsg.), How Taxes Affect Economic Behavior. Washington DC: The Brookings Institution, 27-72.
- Hechtner, F. (2010), Eine theoretische und empirische Studie über Einkommensteuertarife aus Sicht der Wirtschaftswissenschaft – Progressionswirkungen der synthetischen Einkommensteuer, Schedulenbesteuerung und Vermeidung von Doppelbesteuerungen, Berlin 2010.
- Hechtner, F., J. Hundsdoerfer (2009), Schedulenbesteuerung von Kapitaleinkünften mit der Abgeltungsteuer: Belastungswirkungen und neue Problemfelder. *Steuer und Wirtschaft*, 23-41.
- Hechtner, F., J. Hundsdoerfer, C. Sielaff (2011), Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse. *Zeitschrift für betriebswirtschaftliche Forschung*, 214-239.
- Hicks, J. R. (1935), The Theory of Wages. MacMillan, London.
- Houben, H., R. Maiterth (2008), Optimale Nutzung und Wirkungen von § 34a EStG. *Steuer und Wirtschaft*, 228-237.
- Hundsdoerfer, J., C. Sichtmann (2007), Zur Gewichtung von Steuern in unternehmerischen Investitionskalkülen: Eine Conjoint-Analyse des Entscheidungsverhaltens selbständiger Ärzte. *Betriebswirtschaftliche Forschung und Praxis*, 604-621.
- Hundsdoerfer, J., C. Sielaff, K. Blaufus, D. Kiesewetter, J. Weimann (2010), The Name Game for Contributions – Influence of Labeling and Earmarking on the Perceived Tax Burden. *FACTS Discussion Paper*, FU Berlin, 2010/12.
- Keane, M. (2010), Labor Supply and Taxes: A Survey. *UTS Working Paper*, 160.
- Keller, T. (1968), Merkliche und unmerkliche Steuern. *Finanzarchiv*, 100-120.
- Killingsworth, M. (1983), Labor Supply. Cambridge University Press, Cambridge.

- Kimmel, J., T. J. Kniesner (1998), New Evidence on Labor Supply: Employment vs. Hours Elasticities by Sex and Marital Status. *Journal of Monetary Economics*, 42, 289-301.
- Kirchhof, P. (2011), Bundessteuergesetzbuch – Ein Reformentwurf zur Erneuerung des Steuerrechts. Schriftenreihe des Instituts für Finanz- und Steuerrecht: Forschungsstelle Bundessteuergesetzbuch, C. F. Müller, Heidelberg.
- Koalitionsvertrag zwischen CDU, CSU und FDP (2009), Wachstum. Bildung. Zusammenhalt, 17. Legislaturperiode.
- Lewis, A. (1978), Perceptions of Tax Rates. *British Tax Review*, 6, 358-366.
- Maiterth, R. (2003), Beurteilung unterschiedlicher Reformvorschläge für die Gewerbesteuer aus betriebswirtschaftlicher Sicht. *Die Betriebswirtschaft*, 63, 479-496.
- Mann, H., D. Whitney (1947), On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18, 50-60.
- McCaffery, E. J., J. Baron (2003), The Humpty Dumpty blues: Disaggregation bias in the evaluation of tax systems. *Organizational Behavior and Human Decision Processes*, 91 (2), 230-242.
- Mirrlees, J. A. (1971), An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38 (2), 175–208.
- Neumark, F. (1970), Grundsätze gerechter und ökonomisch rationaler Steuerpolitik. Mohr/Siebeck, Tübingen.
- Plott, C. R. (1987), Dimensions of Parallelism: Some Policy Applications of Experimental Methods, in: Roth, A., E. Roth (Hrsg.), *Laboratory Experimentation in Economics: Six Points of View*. Cambridge University Press, New York, NY, 193-229.
- Ramsey, F. P. (1927), A Contribution to the Theory of Taxation. *Economic Journal*, 37 (145), 47-61.
- Rupert, T. J., A. M. Wright (1998), The Use of Marginal Tax Rates in Decision Making: The Impact of Tax Rate Visibility. *Journal of the American Taxation Association*, 20, 83-99.

- Rupert, T. J., C. M. Fischer (1995), An Empirical Investigation of Taxpayer Awareness of Marginal Tax Rates. *Journal of the American Taxation Association*, 17, 36-59.
- Schmölders, G. (1960), Das Irrationale in der öffentlichen Finanzwirtschaft. Rowohlt, Hamburg.
- Sillamaa, M.-A. (1999), Taxpayer behavior in response to taxation: comment and new experimental evidence. *Journal of Accounting and Public Policy*, 18, 165-177.
- Swenson, C. W. (1988), Taxpayers behavior in response to taxation: An experimental analysis. *Journal of Accounting and Public Policy*, 7, 1-28.
- Tipke, K., J. Lang (2010), Steuerrechtsordnung. 20. Aufl., Verlag Dr. Otto Schmidt, Köln.
- Tversky, A., D. Kahneman (1974), Judgment under Uncertainty: Heuristics and Biases. *Science*, 185 (4157), 1124-1131.
- Wagner, F. (1984), Grundfragen und Entwicklungstendenzen der betriebswirtschaftlichen Steuerplanung. *Betriebswirtschaftliche Forschung und Praxis*, 201-221.
- Wagner, F. (1999), Besteuerung, in: Vahlens Kompendium der Betriebswirtschaftslehre, Bd. 2, 4. Aufl., München, 439-504.
- Wagner, F. (2011), Der große Wurf ins Aus, Das Steuerkonzept Kirchhofs ist unklar, widersprüchlich und teuer. *Süddeutsche Zeitung*, 67. Jg., 158, 12. Juli 2011.
- Wilcoxon, F. (1945), Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1, 80-83.
- Ziliak, J. P., T. J. Kniesner (1999), Estimating Life-Cycle Labor Supply Tax Effects. *Journal of Political Economy*, 107, 326-359.

**Diskussionsbeiträge
des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin**

2011

- 2011/1 NEHER, Frank
Markets Wanted – Expectation Overshooting in Transition
Economics
- 2011/2 KNOLL, Martin / Petra ZLOCZYSTI
The Good Governance Indicators of the Millennium Challenge
Account
Economics
- 2011/3 KAPPLER, Marcus / Helmut REISEN / Moritz SCHULARICK /
Edouard TURKISCH
The Macroeconomic Effects of Large Exchange Rate Appreciations
Economics
- 2011/4 MÜLLER, Kai-Uwe / Viktor STEINER
Beschäftigungswirkungen von Lohnsubventionen und Mindestlöhnen
Economics
- 2011/5 WRAGE, Markus / Anja TUSCHKE / Rudi K. F. BRESSER
The Influence of Social Capital on CEO Dismissal in Germany
Strategic Management
- 2011/6 BLAUFUS, Kay / Sebastian EICHFELDER / Jochen
HUNDSDOERFER
The hidden burden of the income tax
FACTS
- 2011/7 MUCHLINSKI, Elke
Die Rezeption der John Maynard Keynes Manuskripte von 1904 bis
1911
Economics
- 2011/8 FOSSEN, Frank M.
Personal bankruptcy law, wealth and entrepreneurship – Theory and
evidence from the introduction of a „fresh start“
Economics
- 2011/9 CALIENDO, Marco / Frank FOSSEN / Alexander KRITIKOS
Personality characteristics and the decision to become and stay
self-employed
Economics

- 2011/10 BACH, Stefan / Martin BEZNOSKA / Viktor STEINER
A Wealth Tax on the Rich to Bring Down Public Debt?
Economics
- 2011/11 HETSCHKO, Clemens / Andreas KNABE / Ronnie SCHÖB
Changing Identity: Retiring from Unemployment
Economics
- 2011/12 BÖRNER, Lars / Battista SEVERGNINI
Epidemic Trade
Economics
- 2011/13 SIELAFF, Christian
Steuerkomplexität und Arbeitsangebot – Eine experimentelle Analyse
FACTS