



# IW-Report 13/19

## Arbeitsangebotsmodul zum IW-Mikrosimulationsmodell STATS

Dokumentation Version 1.0  
Dr. Maximilian Stockhausen

Köln, 10.04.2019

## Inhaltsverzeichnis

<b>1</b>	<b>Erweiterte Modulstruktur des STATS</b>	<b>3</b>
<b>2</b>	<b>Modellierung von Arbeitsangebotsreaktionen</b>	<b>5</b>
2.1	Untersuchungsgruppe	5
2.2	Wahl der Arbeitszeitkategorien ( <i>choice set</i> )	6
2.3	Spezifikation der deterministischen Nutzenfunktion	6
2.4	Heterogenität in den Präferenzen	7
2.5	Imputation fehlender Bruttostundenlöhne	8
<b>3</b>	<b>Berechnung der Arbeitsangebotsreaktionen</b>	<b>10</b>
<b>4</b>	<b>Zusammenfassung</b>	<b>11</b>
<b>5</b>	<b>Anhang</b>	<b>13</b>
	<b>Literatur</b>	<b>21</b>

**JEL-Klassifikation:**

C63: Rechentechniken, Simulationen

C81: Methoden zur Sammlung, Schätzung und Organisation mikroökonomischer Daten

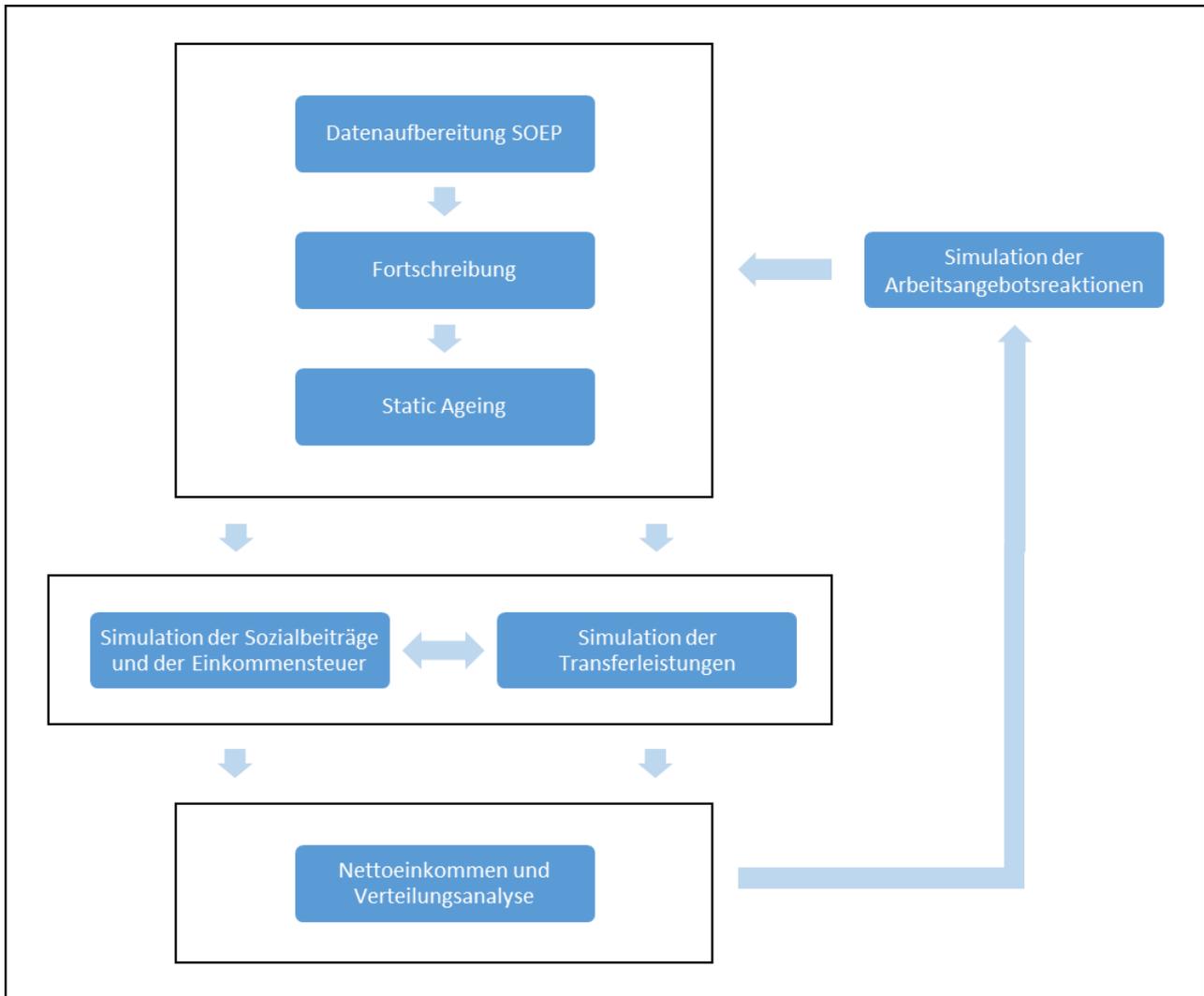
Die Entwicklung und Dokumentation des Arbeitsangebotsmoduls zum IW-Mikrosimulationsmodell STATS ist Bestandteil des von der Carl Deilmann-Stiftung geförderten Projekts „Familienbesteuerung in Deutschland – normative Grundlagen und empirische Fundierungen“.

## 1 Erweiterte Modulstruktur des STATS

Das IW Köln verfügt mit dem Steuer-, Abgaben- und Transfer-Mikrosimulationsmodell (STATS) über ein Analysetool, mit dessen Hilfe die Auswirkungen von Reformen des deutschen Steuer- und Transfersystems auf die Höhe und Verteilung der Nettoeinkommen deutscher Haushalte sowie die Aufkommenswirkungen für den Staat und die einzelnen Sozialversicherungen evaluiert werden können. Das STATS ist modular aufgebaut. Ein erstes Modul dient der Datenaufbereitung. Datenbasis ist das Sozio-ökonomische Panel (SOEP). Das SOEP ist ein repräsentativer Befragungsdatensatz, der seit 1984 im Auftrag des Deutschen Instituts für Wirtschaftsforschung Berlin (DIW) jährlich erhoben wird. Er erfasst alle notwendigen sozio-ökonomischen Indikatoren auf Personen- und Haushaltsebene, darunter verschiedene Einkommensarten, die für die Mikrosimulation erforderlich sind (Wagner et al., 2007). Das zweite und dritte Modul erlauben die Fortschreibung der Daten auf ein beliebiges Analysejahr sowie die Anpassung der Hochrechnungsfaktoren der zugrundeliegenden Beobachtungseinheiten an veränderte Randverteilungen der Bevölkerung. In weiteren Modulen werden anschließend die Einkommensteuer, Sozialbeiträge und Sozialtransfers simuliert. Dabei können Sozialbeiträge und Einkommensteuer auf der einen Seite und Sozialtransfers auf der anderen Seite getrennt voneinander oder simultan berechnet werden. Mit der simulierten Abgabenlast und den Sozialtransfers lassen sich die Nettoeinkommen der Bevölkerung für den Status quo, aber auch für unterschiedliche Reformszenarien berechnen sowie die Aufkommenswirkungen auf aggregierter Ebene kontrollieren. Dabei können die Ergebnisse der statischen Analyse sowohl auf Haushaltsebene als auch auf Personenebene ausgewertet werden (zur Dokumentation von STATS siehe Beznoska, 2016).

Bislang erlaubt die Arbeit mit STATS eine statische Modellierung, das heißt mögliche Anpassungsreaktionen der Wirtschaftssubjekte auf veränderte Arbeitsangebotsanreize, die beispielsweise aus Variationen des Abgaben- und Transferrechts resultieren, können bislang nicht abgebildet werden. Mit der Erweiterung des STATS um ein Modul zur Simulation von Arbeitsangebotsreaktionen können diese Verhaltensänderungen aufgrund veränderter Arbeitsanreize nun modelliert werden, um sogenannte Zweitrundeneffekte zu simulieren. Durch die Integration des Arbeitsangebotsmoduls erfolgt eine wesentliche Weiterentwicklung der bestehenden Modulstruktur: Das STATS wird dynamisiert. Wie sich das Arbeitsangebotsmodul in die Architektur von STATS schematisch einfügt, zeigt Abbildung 1-1. Dabei wird deutlich, dass zur Berechnung der Arbeitsangebotseffekte auf die bereits bestehenden Module zur Berechnung der Haushaltsnettoeinkommen für die unterschiedlichen Arbeitszeitkonstellationen zurückgegriffen wird. Die exakte und in sich konsistente Berechnung der Nettoeinkommen ist dabei eine wesentliche Herausforderung bei der Modellierung und durch einen hohen Komplexitätsgrad gekennzeichnet. Aufbau und Modellannahmen des Erweiterungsmoduls werden in den folgenden Kapiteln beschrieben.

Abbildung 1-1: Erweiterte Modulstruktur des STATS



Eigene Darstellung

## 2 Modellierung von Arbeitsangebotsreaktionen

Zur Modellierung von Arbeitsangebotsreaktionen werden in der einschlägigen Literatur strukturelle Arbeitsangebotsmodelle verwendet (Creedy/Kalb, 2006). All diesen Modellen liegt die Annahme zugrunde, dass Haushalte und die darin befindlichen erwachsenen Personen ihre gemeinsame Nutzenfunktion maximieren. Der Nutzen des Haushalts hängt dabei vom gewählten Konsum und der Arbeitszeit (respektive Freizeit) der Haushaltsmitglieder ab. In den meisten Fällen wird die Arbeitsangebotsentscheidung als ein diskretes Problem modelliert, wobei die Haushaltsmitglieder bei gegebenen Bruttostundenlöhnen eine nutzenoptimale Entscheidung über ihre Wochenarbeitszeit und die damit verbundenen Konsummöglichkeiten treffen. Der Vorteil der Modellierung diskreter Arbeitszeitkategorien liegt darin, dass sich die Nettohaushaltseinkommen exakt bestimmen lassen, keine restriktiven Annahmen über die Form der Budgetgeraden eines Haushalts getroffen werden müssen, und sich die gemeinsame Arbeitsangebotsentscheidung von Paarhaushalten in konsistenter Weise modellieren lässt. Dafür müssen jedoch eine Reihe weiterer Entscheidungen und Modellannahmen getroffen werden, und zwar über:

- die Untersuchungsgruppe,
- die Anzahl der Arbeitszeitkategorien (choice set),
- die Spezifikation der Nutzenfunktion der Haushalte,
- die Berücksichtigung nicht beobachtbarer Heterogenität in den Präferenzen der Haushalte und
- die Imputation fehlender Bruttostundenlöhne.

Die gewählten Spezifikationen des im STATS genutzten Arbeitsangebotsmodells folgen im Wesentlichen älteren Arbeiten von van Soest (1995) und Aaberge et al. (1995) und berücksichtigen neuere Erkenntnisse aus Haan (2006), Aaberge et al. (2009), Aaberge/Colombino (2014, 2018), Löffler et al. (2014) und Jessen et al. (2017).

### 2.1 Untersuchungsgruppe

Die Simulation von Arbeitsangebotsänderungen beschränkt sich auf Haushaltsmitglieder, die ihr Arbeitsangebot infolge veränderter Arbeitsanreize anpassen und damit hinsichtlich ihres Arbeitsangebots flexibel reagieren können. Dazu zählen alle Personen im erwerbsfähigen Alter zwischen 20 und 65 Jahren, die zum Befragungszeitpunkt Arbeiter, Angestellte oder Arbeitslose sind. Ausgeschlossen sind alle Personen, die sich in einer Ausbildung, im Studium, im Wehr- und Zivildienst, in Elternzeit oder in Mutterschutz befinden, die (Früh-)Rentner sind oder eine Behinderung aufweisen, die ihnen eine reguläre Teilnahme am Arbeitsmarkt unmöglich machen. Beamte und Selbständige bleiben ebenso unberücksichtigt, da sie mit anderen Arbeitsmarkt-

und Beschäftigungsbedingungen konfrontiert sind als abhängig Beschäftigte und anderen Anreizstrukturen unterliegen.

Darüber hinaus wird berücksichtigt, ob die Personen in einem Paarhaushalt leben oder nicht. In diesem Zusammenhang werden fünf Gruppen für die Analyse gebildet: Die erste Gruppe umfasst alle Haushalte, in denen beide verheiratete oder zusammenlebende Partner hinsichtlich ihres Arbeitsangebots flexibel sind. Die zweite Gruppe besteht aus allen flexiblen Männern, deren (Ehe)Partnerin unflexibel ist, das heißt, die nicht auf veränderte Arbeitsanreize reagiert oder dem Arbeitsmarkt nicht zur Verfügung steht. Die dritte Gruppe umfasst alle flexiblen Frauen mit einem unflexiblen (Ehe)Partner. Die vierte Gruppe bilden alle flexiblen Single-Männer. Die fünfte Gruppe besteht aus allen flexiblen Single-Frauen. Zu den Gruppen vier und fünf gehören auch alle Alleinerziehenden. Für jede der fünf Gruppen werden die Arbeitsangebotseffekte getrennt voneinander berechnet, um gruppenspezifische Unterschiede in den Präferenzen und Verhaltensreaktionen bestmöglich abbilden zu können, was durch die alleinige Verwendung von Kontrollvariablen (*taste shifters*) möglicherweise nur unzureichend gelingen kann.

## 2.2 Wahl der Arbeitszeitkategorien (*choice set*)

Männer können im Modell zwischen vier diskreten Arbeitszeitkategorien auswählen: 0, 20, 40 und 50 Wochenstunden, während Frauen zwischen sechs Kategorien auswählen können: 0, 10, 20, 30, 40 und 50 Wochenstunden. Diese Unterscheidung ergibt sich aus der Beobachtung, dass immer noch nur sehr wenige Männer in Teilzeit arbeiten. Bei Frauen ist dies hingegen deutlich öfter zu beobachten mit mehr zeitlicher Variation in diesem Bereich. Diesem Unterschied wird durch die unterschiedliche Zahl der Arbeitszeitkategorien Rechnung getragen. Es wird außerdem angenommen, dass jedes Paar/Individuum aus demselben Set an Arbeitszeitkategorien seiner Gruppe/seines Geschlechts auswählen kann. Bei Paaren ergeben sich dadurch 24 mögliche Kombinationen der Arbeitszeitkategorien. Die Auswahl erfolgt auch in Anlehnung an andere, etablierte Mikrosimulationsmodelle. Grundsätzlich ist eine flexiblere Spezifikation der Arbeitszeitkategorien zu empfehlen, das heißt ein *choice set* mit einer größeren Zahl an möglichen Kategorien, zwischen denen die flexiblen Individuen auswählen können. Auch in anderen Arbeiten richtet sich die Wahl der Arbeitszeitkategorien nach der Verteilung der tatsächlich beobachteten Arbeitszeiten wie beispielsweise in Jessen et al. (2017). Die Unterschiede in den Ergebnissen bei Wahl anderer Arbeitszeitkategorien sind jedoch in aller Regel marginal und es zeigt sich, dass der *model fit* nur wenig von der spezifischen Wahl der Arbeitszeitkategorien abhängt, aber durchaus einen Effekt auf die Güte der *out-of-sample* Schätzungen haben kann (Aaberge et al., 2009).

## 2.3 Spezifikation der deterministischen Nutzenfunktion

Diskrete Arbeitsangebotsmodelle gehen weiter von der neo-klassischen Annahme aus, dass Individuen rational handeln und ihren Nutzen maximieren. Paare maximieren eine gemeinsame Nutzenfunktion. Der Nutzen selbst bestimmt sich aus der Entscheidung zwischen Konsum und

Freizeit. Das Konsumniveau wird dabei im Wesentlichen indirekt durch das Arbeitseinkommen bestimmt (sowie Einkommen aus anderen Quellen), sodass ein Individuum bei gegebenem Bruttolohnsatz allein seine Arbeitszeit optimal auswählen muss. Dadurch bestimmt es gleichzeitig über seine Freizeit. Im diskreten Fall kann ein Individuum zwischen fest definierten Arbeitszeitkategorien und der damit verbundenen Freizeit wählen. Für einen Paarhaushalt  $i$  ergibt sich der (gemeinsame) Nutzen  $V$  in der jeweiligen Arbeitszeitkategorie  $j$  als

$$V_{ij} = U(y_{ij}, l_{ij}^f, l_{ij}^m) + \varepsilon_{ij}. \quad (1)$$

Dabei steht  $y$  für das (gemeinsame) Nettoeinkommen des Haushalts und  $l$  für die wöchentliche Freizeit gemäß Arbeitszeitkategorie. Es wird angenommen, dass die maximal mögliche Freizeit pro Woche 80 Stunden für jede Person beträgt. Der Index  $m$  steht für Männer und  $f$  für Frauen und  $\varepsilon$  ist ein Fehlerterm. Für einen Single-Haushalt entfällt entsprechend einer der beiden Freizeitterme. Darüber hinaus erlauben diskrete Arbeitsangebotsmodelle die Anwendung verschiedener Spezifikationen der Nutzenfunktionen. Die häufigsten Spezifikationen sind die *translog*, die *quadratische* oder die *Box-Cox* transformierte Nutzenfunktion (Keane/Moffitt, 1998; Creedy/Kalb, 2006). Im Arbeitsangebotsmodul des STATS wird eine *translog*-Spezifikation zur Modellierung des deterministischen Teils der individuellen Nutzenfunktionen verwendet. Für einen Paarhaushalt, der sein Arbeitsangebot simultan bestimmt und seine gemeinsame Nutzenfunktion maximiert, ergibt sich somit die folgende deterministische Nutzenfunktion:

$$U_{ij} = \beta_1 \ln(y_{ij}) + \beta_2 \ln(l_{ij}^m) + \beta_3 \ln(l_{ij}^f) + \beta_4 \ln(y_{ij})^2 + \beta_5 \ln(l_{ij}^m)^2 + \beta_6 \ln(l_{ij}^f)^2 \quad (2)$$

$$+ \beta_7 \ln(y_{ij}) \ln(l_{ij}^m) + \beta_8 \ln(y_{ij}) \ln(l_{ij}^f) + \beta_9 \ln(l_{ij}^f) \ln(l_{ij}^m).$$

Die Notation ist dieselbe wie zuvor. Für einen Single-Haushalt verkürzt sich die Funktion entsprechend um einen der beiden geschlechtsspezifischen Freizeitterme und die jeweiligen Kreuzterme.

## 2.4 Heterogenität in den Präferenzen

Beobachtbare Unterschiede zwischen den Haushalten, die sich auf die Präferenzen für Arbeit, respektive Freizeit und Einkommen auswirken können, lassen sich durch eine Erweiterung von Gleichung (2) ebenfalls modellieren. Dazu wird eine Reihe individueller und haushaltspezifischer Charakteristika (in  $X_1, X_2$  und  $X_3$  enthalten) mit den Parametern für Arbeit und Einkommen der ersten Ordnung interagiert und als Kontrollvariablen dem Modell hinzugefügt:

$$\begin{aligned} \beta_1 &= \alpha_0^y + X_1' \alpha_1^y \\ \beta_2 &= \alpha_0^m + X_2' \alpha_1^m \\ \beta_3 &= \alpha_0^f + X_3' \alpha_1^f \end{aligned} \quad (3)$$

Zu den Kontrollvariablen gehören unter anderem: Die Anzahl der im Haushalt lebenden Kinder zwischen null und sechs Jahren, das Alter und das Alter im Quadrat der jeweiligen Personen, ein regionaler Dummy, der anzeigt, ob ein Haushalt in West- oder Ostdeutschland lebt oder ein Dummy, ob ein Migrationshintergrund vorliegt. Zudem werden potenzielle Fixkosten einer Teilzeitbeschäftigung von  $\leq 30$  Wochenstunden berücksichtigt, die zumeist mit einem negativen Nutzen für die Betroffenen einhergehen. Zu diesen Kosten können beispielsweise erhöhte individuelle Suchkosten zählen oder der Umstand, dass mehr oder weniger als die gewünschten Stunden gearbeitet werden muss (Arbeitsmarkttrigiditäten). Die genauen Modellspezifikationen für die einzelnen Gruppen ergeben sich aus Tabelle 5-2 bis Tabelle 5-6 im Anhang.

Unter Berücksichtigung dieser beobachtbaren Heterogenität und der Annahme, dass die Fehlerterme unabhängig und identisch über die Arbeitszeitkategorien und Haushalte verteilt sind und einer *extreme-value type I* Verteilung folgen, lässt sich die Auswahlwahrscheinlichkeit der Arbeitszeitkategorie  $k$  durch Haushalt  $i$  mithilfe eines konditionalen Logit Modells nach McFadden (1974) schätzen:

$$P_{ik} = \Pr(V_{ik} > V_{ij}, \forall j = 1, \dots, J) = \frac{\exp(U_{ik})}{\sum_{j=1}^J \exp(U_{ij})}, \quad k \in J. \quad (4)$$

Die Arbeitszeitkategorie  $k$  wird dann von einer Person beziehungsweise einem Haushalt gewählt, wenn der Nutzen aus dieser Kategorie höher ist als in allen anderen Fällen. Für unbeobachtete Heterogenität zwischen den Haushalten wird in der derzeitigen Modellversion nicht weiter kontrolliert. Diese Erweiterung wird für die Zukunft angestrebt. Allerdings können Jessen et al. (2017) zeigen, dass dies nur geringe Auswirkungen auf die Qualität der Schätzergebnisse hat und verweisen auf ähnliche Resultate in Haan (2006).

## 2.5 Imputation fehlender Bruttostundenlöhne

Will man die Arbeitsangebotsreaktionen von flexiblen Personen bestimmen, die zum gegebenen Befragungszeitpunkt dem Arbeitsmarkt nicht zur Verfügung stehen, weil sie beispielsweise temporär erwerblos sind oder Kinder zu betreuen haben, müssen die fehlenden Bruttostundenlöhne geschätzt werden. Das gleiche gilt für Personen, die aus anderen Gründen über fehlende Angaben verfügen (*item non-response*). Dazu wird ein zweistufiges Schätzverfahren nach Heckman (1979) angewendet: In einem ersten Schritt wird die Wahrscheinlichkeit bestimmt, am Arbeitsmarkt teilzunehmen (Selektionsregression), und in einem zweiten Schritt wird die Höhe der Bruttostundenlöhne geschätzt. Dabei wird berücksichtigt, dass ein möglicher Zusammenhang zwischen der Arbeitsmarktteilnahme und der Bruttostundenlohnhöhe besteht, der auf nicht beobachtbare Eigenschaften oder andere Selektionseffekte der Personen zurückgeht. Als Selektionsvariablen werden genutzt: der Familienstand, die Anzahl von Kindern im Haushalt zwischen null und drei Jahren und zwischen vier und vierzehn Jahren, der selbsteingeschätzte Gesundheitszustand, eine Dummy-Variable für das Vorliegen einer Erwerbs- oder Schwerbehinderung und eine Variable, die die Summe sonstigen Einkommens (Kapitaleinkommen, private und öffentliche Transfers, private und öffentliche Renten) des Haushalts umfasst.

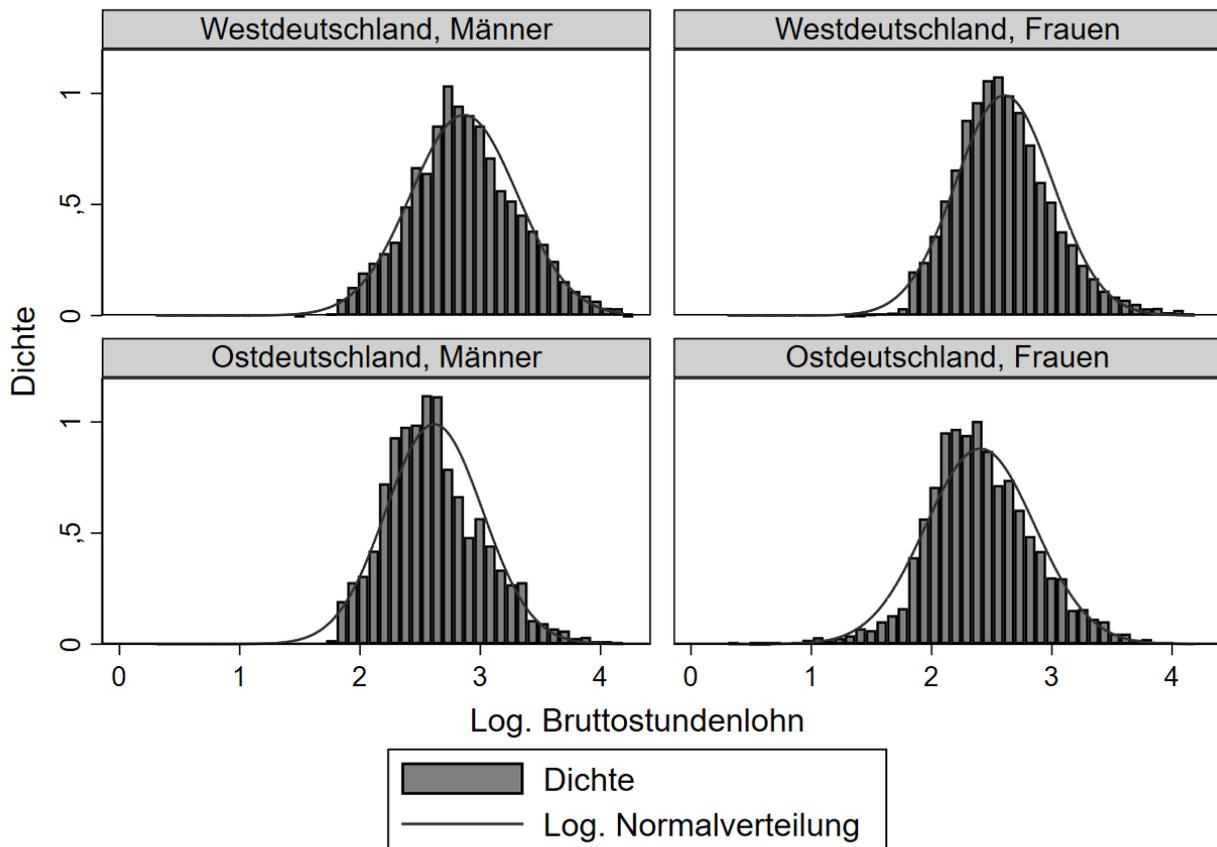
Die Heckman-Schätzung zur Imputation der Bruttostundenlöhne erfolgt zu Beginn der Simulation und greift auf die Informationen von allen Personen im jeweiligen Beobachtungsjahr zurück, die im erwerbsfähigen Alter zwischen 20 und 65 Jahren sind, einer abhängigen Beschäftigung nachgehen oder die zum Zeitpunkt der Beobachtung nicht erwerbstätig beziehungsweise arbeitslos sind. Zu den berücksichtigten Erwerbstätigkeiten zählen sowohl Vollzeit- und Teilzeitbeschäftigungen als auch geringfügige Beschäftigungsverhältnisse. Ausgeschlossen sind Auszubildende, Rentner, Wehr- und Zivildienstleistende, Volontäre und Praktikanten, Beamte und Selbständige aller Formen. Die Schätzungen werden zudem für West- und Ostdeutschland und für Männer und Frauen getrennt voneinander durchgeführt. Die Schätzergebnisse des Heckman-Modells werden am Beispiel westdeutscher Single Männer in Tabelle 5-1 des Anhangs dargestellt. Selektionseffekte bestehen nachweislich für west- und ostdeutsche Frauen. Für Männer hingegen nicht.

Da im SOEP die individuellen Bruttostundenlöhne nicht direkt abgefragt werden, werden diese hier als Quotient aus dem jährlichen Arbeitslohn und den jährlich geleisteten Arbeitsstunden ermittelt. Dabei kommt es insbesondere an den Rändern der Lohnverteilung zu unplausibel geringen und hohen Bruttostundenlöhnen, sodass im Vorfeld die unteren 10 Prozent ( $< 6$  Euro/h) und das oberste Prozent ( $> 66$  Euro/h) auf *Missing* gesetzt werden und somit von der Heckman-Schätzung ausgeschlossen sind.

Mithilfe der Modellparameter des Heckman-Modells werden dann die Bruttostundenlöhne für alle Personen im SOEP geschätzt. Die geschätzten Bruttostundenlöhne werden jedoch nur bei denjenigen Personen verwendet, bei denen die Information zum Bruttostundenlohn fehlt. Alle anderen Personen behalten ihre beobachteten Bruttostundenlöhne bei. Ein Vorteil dieses Vorgehens ist die Erhaltung der ursprünglichen Varianz der Bruttostundenlöhne in der Stichprobe. Alternativ könnten die geschätzten Bruttostundenlöhne für alle Personen verwendet werden. Dies würde zu einer Verringerung der Varianz in den Daten führen und zu einer späteren Überschätzung der Arbeitsangebotselastizitäten (Löffler et al., 2014).

Die resultierende Verteilung der Bruttostundenlöhne nach Region und Geschlecht kann Abbildung 2-1 entnommen werden. Für westdeutsche Männer und Frauen folgen die logarithmierten Bruttolöhne der logarithmierten Normalverteilung, während die Verteilung für ostdeutsche Männer und Frauen leicht rechtsschief ist, aber auch in diesem Fall annähernd der logarithmierten Normalverteilung folgt. Abschließend sei auf die Annahme hingewiesen, dass die ermittelten Bruttostundenlöhne bei allen anderen Entscheidungen über das zu wählende Arbeitsangebot exogen gegeben sind und nicht von der jeweiligen Arbeitszeitkategorie und damit verbundenen Präferenzen abhängen.

**Abbildung 2-1: Verteilung der Bruttostundenlöhne nach Region und Geschlecht**



Quellen: SOEP v32.1, eigene Berechnungen

### 3 Berechnung der Arbeitsangebotsreaktionen

Alle Veränderungen der Nettoeinkommen eines Haushaltes, ob sie durch eine exogene Erhöhung der Bruttostundenlöhne oder durch andere Veränderungen im Steuer- und Transfersystem verursacht werden, wirken sich unmittelbar auf die Auswahlwahrscheinlichkeiten der einzelnen Arbeitszeitkategorien aus. Unter Anwendung der sogenannten *probability method* kann die unkompenzierte Arbeitsangebotselastizität der Haushalte als Quotient aus der prozentualen Veränderung des erwarteten Arbeitsangebots und der prozentualen Veränderung der jeweiligen Einkommenskomponente berechnet werden (Creedy/Kalb, 2006).

Die sich aus einer jeweils zehnpromzentigen Erhöhung der Bruttostundenlöhne von Männern und Frauen ergebenden unkompenzierten Arbeitsangebotselastizitäten sind in Tabelle 3-1 abgetragen. Es werden sowohl die durchschnittlichen als auch die aggregierten Elastizitäten ausgewiesen. Die geschätzten unkompenzierten Arbeitsangebotselastizitäten liegen dabei im Wertebereich zwischen 0,02 und 0,08. Grundsätzlich fallen die Anpassungsreaktionen bei Frauen größer

aus als bei Männern. Für beide Geschlechter zeigt sich zudem, dass das Arbeitsangebot von Single-Haushalten stärker auf eine zehnpromtente Lohnerhöhung reagiert als bei Paarhaushalten. Dabei fallen die durchschnittlichen Lohnelastizitäten tendenziell etwas größer aus als die aggregierten. Des Weiteren führt eine zehnpromtente Lohnerhöhung bei Männern mit flexiblen Partnern zu einer Reduzierung des Arbeitsangebots der jeweiligen Partnerin. Männer reagieren hingegen auf eine zehnpromtente Lohnerhöhung ihrer Partnerin praktisch nicht. Insgesamt entsprechen die erzielten Angebotsreaktionen den theoretischen Vorüberlegungen und liegen im Bereich vergleichbarer neuerer Studien (Jessen et al., 2017).

### Tabelle 3-1: Unkompensierte Arbeitsangebotselastizitäten

Effekte einer Erhöhung der Bruttostundenlöhne von Männern und Frauen um jeweils 10 Prozent; durchschnittliche und aggregierte Bruttostundenlohnelastizitäten

Gruppe	Durchschnittliche		Aggregierte	
	Lohnelastizität	Lohnelastizität des Partners	Lohnelastizität	Lohnelastizität des Partners
Paar – beide flexibel				
Mann	0,021	-0,018	0,020	-0,013
Frau	0,049	0,001	0,044	0,001
Paar – einer flexibel				
Mann	0,037	/	0,037	/
Frau	0,043	/	0,038	/
Singles – flexibel				
Mann	0,048	/	0,049	/
Frau	0,076	/	0,068	/

Quellen: SOEP v32.1, eigene Berechnungen

## 4 Zusammenfassung

Mit dem dargestellten Arbeitsangebotsmodul wird das Steuer-, Abgaben- und Transfer-Mikrosimulationsmodell des IW Köln um eine tragende Komponente erweitert, die die Simulation von Zweitrundeneffekten in Folge von Änderungen des deutschen Steuer- und Transfersystems auf unterschiedliche Zielgrößen ermöglicht. Am Beispiel einer (exogenen) Erhöhung der Bruttostundenlöhne von Männern und Frauen um jeweils 10 Prozent ließ sich zeigen, wie sich das Arbeitsangebot von unterschiedlichen Haushaltstypen anpasst. Die geschätzten unkompenzierten Arbeitsangebotselastizitäten liegen dabei zwischen 0,02 und 0,08. Die Arbeitsangebotseffekte sind dabei für Single-Haushalte tendenziell größer als für Paarhaushalte und für Frauen größer als für Männer.

Dabei sind dem Modell jedoch auch Grenzen gesetzt: So bezieht sich die Simulation von Arbeitsangebotseffekten ausschließlich auf abhängig Beschäftigte im erwerbsfähigen Alter zwischen 20 und 65 Jahren, die auf veränderte Arbeitsanreize auf dem Arbeitsmarkt reagieren können. Das Verhalten von Beamten, Rentnern oder Selbständigen wird aufgrund von andersartigen Arbeitsmarkt- und Beschäftigungsbedingungen in der derzeitigen Modellversion nicht abgebildet. Dies steht ebenfalls in Zusammenhang mit unterschiedlichen Präferenzen dieser Gruppen, die nicht oder nur schwer beobachtet werden können.

Zudem unterliegt die Modellschätzung zum Teil strikten Annahme bezüglich des nutzenoptimalen Verhaltens der Haushalte und der darin lebenden Individuen: So kann beispielsweise die Annahme diskutiert werden, ob Paarhaushalte stets eine gemeinsame Nutzenfunktion maximieren – wie es angenommen wird – oder ob die Individuen nicht ihren eigenen Nutzen unter Berücksichtigung der Entscheidungen des Partners optimieren. Trotz dieser Einschränkungen stellen diskrete Arbeitsangebotsmodule bisweilen die beste und flexibelste Möglichkeit zur Modellierung von Arbeitsangebotsentscheidungen dar und sind ein wertvolles Instrument zur Evaluierung unterschiedlicher sozialpolitischer Reformvorhaben.

## 5 Anhang

**Tabelle 5-1: Ergebnisse der zweistufigen Heckman-Bruttostundenlohnregression am Beispiel westdeutscher Männer**

Variable	Koeffizient	Standardfehler	p-Wert
<b>Lohnregression (N=4084)</b>			
Alter	0,0235	0,0073	0,0014
Alter <sup>2</sup>	-0,0003	0,0001	0,0001
Arbeitsmarkterfahrung in Teilzeit	-0,0301	0,0049	0,0000
(Arbeitsmarkterfahrung in Teilzeit) <sup>2</sup>	0,0010	0,0003	0,0001
Arbeitsmarkterfahrung in Vollzeit	0,0272	0,0040	0,0000
(Arbeitsmarkterfahrung in Vollzeit) <sup>2</sup>	-0,0003	0,0001	0,0004
Migrationshintergrund (Referenz: keiner)			
Direkter	-0,1368	0,0159	0,0000
Indirekter	-0,0101	0,0199	0,6115
Bildungsjahre	0,0895	0,0033	0,0000
Konstante	1,0943	0,1432	0,0000
<b>Selektionsregression (N=650)</b>			
Alter	0,0515	0,0242	0,0333
Alter <sup>2</sup>	-0,0016	0,0003	0,0000
Arbeitsmarkterfahrung in Teilzeit	0,0659	0,0205	0,0013
(Arbeitsmarkterfahrung in Teilzeit) <sup>2</sup>	0,0002	0,0011	0,8741
Arbeitsmarkterfahrung in Vollzeit	0,1152	0,0120	0,0000
(Arbeitsmarkterfahrung in Vollzeit) <sup>2</sup>	-0,0005	0,0003	0,0464
Migrationshintergrund (Referenz: keiner)			
Direkter	-0,5721	0,0626	0,0000
Indirekter	-0,1707	0,0874	0,0509
Bildungsjahre	0,1352	0,0117	0,0000
Selbsteingeschätzter Gesundheitszustand (Referenz: sehr gut)			
Gut	0,1256	0,0875	0,1509
Zufriedenstellend	0,0342	0,0943	0,7170
Schlecht	-0,3410	0,1084	0,0017
Sehr schlecht	-0,8744	0,1617	0,0000
Familienstand (Referenz: verheiratet/zusammenlebend)			
Single	-0,4132	0,0847	0,0000
Verwitwet	-0,2407	0,3337	0,4706
Geschieden	-0,3641	0,1049	0,0005
Getrennt lebend	-0,4786	0,1852	0,0098

Erwerbs- oder Schwerbehinderung	-0,2647	0,1004	0,0083
Kind im Alter 0-3 Jahre	-0,2688	0,0820	0,0010
Kind im Alter 4-13 Jahre	0,0580	0,0675	0,3903
Sonstiges Einkommen	-0,0002	0,0000	0,0000
Konstante	-0,8541	0,4692	0,0687
Lambda	0,0199	0,0444	0,6548
Beobachtungen insgesamt	4734		

Quellen: SOEP v32.1, eigene Berechnungen

**Tabelle 5-2: Schätzergebnisse des diskreten Arbeitsangebotsmodells: Paar – beide flexibel**

Abhängige Variable: (Gewählte) Arbeitszeitkategorie

Variable	Koeffizient	Standardfehler	p-Wert
Log. Nettoeinkommen	-2,498	2,752	0,364
(Log. Nettoeinkommen) <sup>2</sup>	-0,055	0,050	0,278
Log. Freizeit Mann	71,091	7,470	0,000
(Log. Freizeit Mann) <sup>2</sup>	-9,916	0,544	0,000
Log. Freizeit Frau	111,659	11,539	0,000
(Log. Freizeit Frau) <sup>2</sup>	-15,991	1,287	0,000
Log. Nettoeinkommen x log. Freizeit Mann	-0,193	0,412	0,639
Log. Nettoeinkommen x log. Freizeit Frau	-0,145	0,392	0,712
Log. Freizeit Mann x log. Freizeit Frau	3,953	0,845	0,000
Alter Mann x log. Nettoeinkommen	-0,061	0,098	0,531
(Alter Mann) <sup>2</sup> x log. Nettoeinkommen	0,001	0,001	0,217
Alter Frau x log. Nettoeinkommen	0,306	0,106	0,004
(Alter Frau) <sup>2</sup> x log. Nettoeinkommen	-0,004	0,001	0,002
Ostdeutschland x log. Nettoeinkommen	-1,111	0,325	0,001
Migrationshintergrund Frau x log. Nettoeinkommen	-0,973	0,261	0,000
Migrationshintergrund Frau x log. Freizeit Frau	0,811	0,436	0,063
Alter Frau x log. Freizeit Frau	0,083	0,134	0,535
(Alter Frau) <sup>2</sup> x log. Freizeit Frau	-0,001	0,002	0,704
Ostdeutsche Frau x log. Freizeit Frau	-10,216	5,287	0,053
Kind 0-5 x log. Freizeit Frau	2,756	0,367	0,000
Teilzeit Frau (10h)	-0,907	0,136	0,000
Teilzeit Frau (20h)	-1,134	0,150	0,000
Teilzeit Frau (30h)	-1,505	0,161	0,000
Migrationshintergrund Mann x log. Nettoeinkommen	-0,110	0,333	0,741
Migrationshintergrund Mann x	1,975	0,575	0,001

log. Freizeit Mann			
Alter Mann x log. Freizeit Mann	-0,424	0,156	0,007
(Alter Mann) <sup>2</sup> x log. Freizeit Mann	0,005	0,002	0,008
Ostdeutscher Mann x log. Freizeit Mann	-8,301	5,469	0,129
Teilzeit Mann (20h)	-1,573	0,161	0,000
Migrationshintergrund Mann x log. Freizeit Mann x log. Freizeit Frau	-0,140	0,114	0,217
Ostdeutschland x log. Freizeit Mann x log. Freizeit Frau	2,067	1,399	0,139
#Beobachtungen	67.104		
Pseudo R <sup>2</sup>	0,20		

Anmerkung: Gewichtet mit Haushaltsgewichten.

Quellen: SOEP v32.1, eigene Berechnungen

**Tabelle 5-3: Schätzergebnisse des diskreten Arbeitsangebotsmodells: Paar – Mann flexibel**

Abhängige Variable: (Gewählte) Arbeitszeitkategorie

Variable	Koeffizient	Standardfehler	p-Wert
Log. Nettoeinkommen	2,006	4,921	0,684
(Log. Nettoeinkommen) <sup>2</sup>	-0,013	0,091	0,886
Log. Freizeit Mann	75,920	8,690	0,000
(Log. Freizeit Mann) <sup>2</sup>	-8,114	0,715	0,000
Log. Nettoeinkommen x log. Freizeit Mann	-0,470	0,604	0,436
Alter Mann x log. Nettoeinkommen	-0,013	0,239	0,958
(Alter Mann) <sup>2</sup> x log. Nettoeinkommen	0,001	0,003	0,833
Ostdeutschland Mann x log. Nettoeinkommen	-0,622	0,442	0,160
Migrationshintergrund Mann x log. Nettoeinkommen	0,395	0,418	0,345
Migrationshintergrund Mann x log. Freizeit Mann	2,502	0,545	0,000
Alter Mann x log. Freizeit Mann	-0,473	0,234	0,043
(Alter Mann) <sup>2</sup> x log. Freizeit Mann	0,005	0,003	0,061
Ostdeutscher Mann x log. Freizeit Mann	-0,025	0,816	0,975
Teilzeit Mann (20h)	-2,095	0,235	0,000
#Beobachtungen	5.372		
Pseudo R <sup>2</sup>	0,35		

Anmerkung: Gewichtet mit Haushaltsgewichten.

Quellen: SOEP v32.1, eigene Berechnungen

**Tabelle 5-4: Schätzergebnisse des diskreten Arbeitsangebotsmodells: Paar – Frau flexibel**

Abhängige Variable: (Gewählte) Arbeitszeitkategorie

Variable	Koeffizient	Standardfehler	p-Wert
Log. Nettoeinkommen	-8,489	3,397	0,012
(Log. Nettoeinkommen) <sup>2</sup>	0,191	0,064	0,003
Log. Freizeit Frau	87,209	15,151	0,000
(Log. Freizeit Frau) <sup>2</sup>	-11,291	1,804	0,000
Log. Nettoeinkommen x log. Freizeit Frau	0,602	0,529	0,255
Alter Frau x log. Nettoeinkommen	0,177	0,091	0,050
(Alter Frau) <sup>2</sup> x log. Nettoeinkommen	-0,001	0,001	0,132
Ostdeutschland Frau x log. Nettoeinkommen	-0,086	0,301	0,776
Migrationshintergrund Frau x log. Nettoeinkommen	-0,455	0,362	0,209
Migrationshintergrund Frau x log. Freizeit Frau	1,007	0,508	0,048
Alter Frau x log. Freizeit Frau	-0,151	0,170	0,376
(Alter Mann) <sup>2</sup> x log. Freizeit Frau	0,003	0,002	0,207
Ostdeutsche Frau x log. Freizeit Frau	-1,355	0,553	0,014
Kind 0-5 x log. Freizeit Frau	2,516	0,664	0,000
Teilzeit Frau (10h)	-1,027	0,227	0,000
Teilzeit Frau (20h)	-0,817	0,231	0,000
Teilzeit Frau (30h)	-1,355	0,256	0,000
#Beobachtungen	5.064		
Pseudo R <sup>2</sup>	0,11		

Anmerkung: Gewichtet mit Haushaltsgewichten.

Quellen: SOEP v32.1, eigene Berechnungen

**Tabelle 5-5: Schätzergebnisse des diskreten Arbeitsangebotsmodells: Single – Mann**

Abhängige Variable: (Gewählte) Arbeitszeitkategorie

Variable	Koeffizient	Standardfehler	p-Wert
Log. Nettoeinkommen	0,749	1,603	0,640
(Log. Nettoeinkommen) <sup>2</sup>	0,033	0,030	0,267
Log. Freizeit Mann	61,591	6,835	0,000
(Log. Freizeit Mann) <sup>2</sup>	-6,760	0,696	0,000
Log. Nettoeinkommen x log. Freizeit Mann	-0,085	0,314	0,787
Alter Mann x log. Nettoeinkommen	0,001	0,037	0,983
(Alter Mann) <sup>2</sup> x log. Nettoeinkommen	0,000	0,000	0,953
Ostdeutschland Mann x log. Nettoeinkommen	-0,508	0,166	0,002
Migrationshintergrund Mann x log. Nettoeinkommen	-0,683	0,154	0,000
Migrationshintergrund Mann x log. Freizeit Mann	0,623	0,443	0,160
Alter Mann x log. Freizeit Mann	-0,303	0,128	0,018
(Alter Mann) <sup>2</sup> x log. Freizeit Mann	0,003	0,002	0,042
Ostdeutscher Mann x log. Freizeit Mann	-0,489	0,521	0,348
Teilzeit Mann (20h)	-1,780	0,158	0,000
#Beobachtungen	3.412		
Pseudo R <sup>2</sup>	0,21		

Anmerkung: Ungewichtet.

Quellen: SOEP v32.1, eigene Berechnungen

**Tabelle 5-6: Schätzergebnisse des diskreten Arbeitsangebotsmodells: Single – Frau**

Abhängige Variable: (Gewählte) Arbeitszeitkategorie

Variable	Koeffizient	Standardfehler	p-Wert
Log. Nettoeinkommen	-0,540	1,912	0,778
(Log. Nettoeinkommen) <sup>2</sup>	0,088	0,035	0,013
Log. Freizeit Frau	93,566	10,820	0,000
(Log. Freizeit Frau) <sup>2</sup>	-10,983	1,252	0,000
Log. Nettoeinkommen x log. Freizeit Frau	-0,192	0,379	0,612
Alter Frau x log. Nettoeinkommen	0,053	0,051	0,295
(Alter Frau) <sup>2</sup> x log. Nettoeinkommen	-0,001	0,001	0,405
Ostdeutschland Frau x log. Nettoeinkommen	-0,309	0,170	0,069
Migrationshintergrund Frau x log. Nettoeinkommen	-0,547	0,155	0,000
Migrationshintergrund Frau x log. Freizeit Frau	0,874	0,449	0,051
Alter Frau x log. Freizeit Frau	-0,251	0,150	0,093
(Alter Mann) <sup>2</sup> x log. Freizeit Frau	0,003	0,002	0,062
Ostdeutsche Frau x log. Freizeit Frau	0,065	0,437	0,882
Kind 0-5 x log. Freizeit Frau	2,716	0,761	0,000
Teilzeit Frau (10h)	-1,760	0,201	0,000
Teilzeit Frau (20h)	-2,035	0,193	0,000
Teilzeit Frau (30h)	-1,635	0,172	0,000
#Beobachtungen	10.236		
Pseudo R <sup>2</sup>	0,14		

Anmerkung: Gewichtet mit Haushaltsgewichten.

Quellen: SOEP v32.1, eigene Berechnungen

## Literatur

- Aaberge, R. / Colombino, U. / Wennemo, T., 2009, Evaluating Alternative Representations of the Choice Sets in Models of Labor Supply, in: *Journal of Economic Surveys*, 23. Jg., Nr. 3, S. 586–612
- Aaberge, Rolf / Colombino, Ugo, 2014, Labour Supply Models, in: *Contributions to Economic Analysis*, Bd. 293, *Handbook of Microsimulation Modelling*, S. 167–221
- Aaberge, Rolf / Colombino, Ugo, 2018, Structural Labour Supply Models and Microsimulation, in: *IZA Discussion Paper*, Nr. 11562, S. 1–38
- Aaberge, Rolf / Dagsvik, John K. / Strøm, Steinar, 1995, Labor Supply Responses and Welfare Effects of Tax Reforms, in: *The Scandinavian Journal of Economics*, 97. Jg., Nr. 4, S. 635–659
- Beznoska, Martin, 2016, Dokumentation zum Steuer-, Abgaben- und Transfer-Mikrosimulationsmodell des IW Köln (STATS), *IW-Report*, Nr. 27, Köln
- Creedy, John / Kalb, Guyonne, 2006, *Labour supply and microsimulation. The evaluation of tax policy reforms*, Cheltenham
- Haan, Peter, 2006, Much ado about nothing: conditional logit vs. random coefficient models for estimating labour supply elasticities, in: *Applied Economics Letters*, 13. Jg., Nr. 4, S. 251–256
- Heckman, James J., 1979, Sample Selection Bias as a Specification Error, in: *Econometrica*, 47. Jg., Nr. 1, S. 153–161
- Jessen, Robin / Rostam-Afschar, Davud / Steiner, Viktor, 2017, Getting the Poor to Work: Three Welfare-Increasing Reforms for a Busy Germany, in: *FinanzArchiv: Public Finance Analysis*, 73. Jg., Nr. 1, S. 1–41
- Keane, Michael / Moffitt, Robert, 1998, A Structural Model of Multiple Welfare Program Participation and Labor Supply, in: *International Economic Review*, 39. Jg., Nr. 3, S. 553–589
- Löffler, Max / Peichl, Andreas / Siegloch, Sebastian, 2014, Structural Labor Supply Models and Wage Exogeneity, in: *IZA Discussion Paper*, Nr. 8281, S. 1–44
- McFadden, D., 1974, Conditional Logit Analysis of Qualitative Choice Behavior, in: Zarembka, P. (Hrsg.), *Frontiers in Econometrics*, New York, S. 105–142
- van Soest, Arthur, 1995, Structural Models of Family Labor Supply: A Discrete Choice Approach, in: *The Journal of Human Resources*, 30. Jg., Nr. 1, S. 63–88

Wagner, Gert G. / Frick, Joachim R. / Schupp, Jürgen, 2007, The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements, in: Schmollers Jahrbuch, 127. Jg., Nr. 1, S. 139–170