
Swiss Inequality Database (SID) Health: Technischer Appendix.

STAND: 20. SEPTEMBER 2024

Inhaltsverzeichnis

1	Einkommensmodellierung	3
1.1	Einleitung	3
1.2	Modell	3
1.3	Unabhängige Variablen	3
1.4	Prognosestärke	5
1.5	Resultate Regressionsmodell	6
1.6	Definition Einkommensgruppen und deskriptive Statistiken	7
2	Umverteilungskomponenten	8
2.1	Gesamte OKP-Umverteilung	8
	Versicherungstechnisches Ergebnis	8
2.2	Prämienzahlungen	9
2.3	Individuelle Prämienverbilligungen (IPV)	9
2.4	Kantonale Finanzierung der stationären Spitalleistungen	10
3	Berechnung Steuerfinanzierung	11
4	Literaturverzeichnis	13
	Anhang	14

1 Einkommensmodellierung

1.1 Einleitung

Die SID Health analysiert die Umverteilungseffekte über die Einkommensverteilung. Da die Krankenversicherer in der OKP nicht über Einkommensdaten der Versicherten verfügen, müssen die Einkommen modelliert werden. Dies macht die Einkommensmodellierung, anhand derer die Individuen in vier Einkommensgruppen (unterste Einkommen, unterer Mittelstand, oberer Mittelstand und oberste Einkommen) eingeteilt werden, zu einem zentralen Methoden-Baustein der SID Health. Als Ausgangspunkt dient das Einkommensmodell von Schmid et al. (2022), das in einer verfeinerten Form verwendet wird.

1.2 Modell

Das Einkommen wird anhand von Berufsinformationen approximiert, die für einen Teil der CSS-Versicherten verfügbar sind. Hier wurden ausgewählte Berufe als Hoch- oder Niedriglohnberufe klassifiziert. Grundsätzlich wird Personen mit einem klassifizierten Hochlohnberuf der Wert $Y = 1$ und Personen mit einem klassifizierten Niedriglohnberuf der Wert $Y = 0$ zugewiesen. Dem Rest der Individuen wird mithilfe des folgenden logistischen Regressionsmodells ein Y -Wert zugeordnet:

$$\text{Logit}(Y_{1/0}|X_i = x_i) = \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + \dots + x_{ik}\beta_k + \epsilon_i$$

Vor der Approximation der fehlenden Werte mittels des Regressionsmodells werden die Berufsindikatoren auf die Familienmitglieder mit gleichem Wohnsitz erweitert, wobei jeweils auf Familienebene der Höchstwert der Berufsindikatoren zugewiesen wird (d.h. das Familieneinkommen wird als «hoch» klassifiziert, wenn mindestens eine Person in der Familie einen Hochlohnberuf hat). Anschliessend wird pro Familie eine Person zufällig ausgewählt und die restlichen Familienmitglieder werden ausgeschlossen, um die Unabhängigkeit der Untersuchungseinheiten zu garantieren.¹

Zur Schätzung der Hochlohnwahrscheinlichkeit werden die verschiedenen unabhängigen Variablen X verwendet, die im nachfolgenden Abschnitt genauer beschrieben werden.

1.3 Unabhängige Variablen

Altersgruppen: Es wird für die Altersgruppe kontrolliert, der die versicherte Person angehört (Datenquelle CSS). Mit Ausnahme der 19–20-Jährigen werden die Altersgruppen in 5-Jahres-Intervallen definiert. Pensionierte werden in einer Altersgruppe zusammengefasst, weil

¹ Anhand einer Altersanalyse wurde überprüft, ob dieses Vorgehen zu einer systematischen Überschätzung der Einkommen von jungen Erwachsenen führt, indem das durchschnittliche Einkommen der 19–20-Jährigen, mit dem der 31–35-Jährigen verglichen wurde. Dabei zeigt sich, dass die Einkommen der 31–35-Jährigen erwartungsgemäss höher sind, mögliche Verzerrungen nicht stark ins Gewicht fallen.

die Berufsinformationen² im hohen Alter spärlich vorhanden sind und dadurch die Schätzungen verzerrt würden. Die Referenzgruppe ist die Altersgruppe 19-20.³

Franchise: Eine weitere Kontrollvariable stellt die gewählte Franchise in der Grundversicherung dar (Datenquelle CSS).

Einkommensindex: Als Kontrollvariable wird auch die ursprüngliche Einkommensmodellierung aus bisherigen Forschungspublikationen (z.B. Schmid et al. (2022)) verwendet (Datenquelle CSS). Als Basis dieses Indexes dienen dieselben Berufsklassifizierungen, wie für die abhängige Variable dieses Modells. Der Wert ist ein gewichtetes Mittel aus den Berufsklassifizierungen der Nachbarn, wobei das Gewicht mit zunehmender Distanz abnimmt. Der Wertebereich liegt zwischen -1 und $+1$.

IPV-Empfänger: Als weitere Kontrollvariable wird die Information verwendet, ob die versicherte Person individuelle Prämienverbilligungen (IPV) erhält (Datenquelle CSS). Die Variable ist nur für die Jahre ab 2014 verfügbar, da in den vorherigen Jahren die IPV teilweise direkt an die Versicherten ausbezahlt wurden. Deshalb beginnt der gesamte Analysezeitraum erst ab 2014.

Monatsprämie VVG: Es wird für die monatliche Prämienhöhe für Zusatzversicherungsprodukte, die von der versicherten Person bezahlt wird, kontrolliert (Datenquelle CSS).

Median-Einkommen Gemeinde: Das Median-Einkommen der Wohngemeinde des Individuums fungiert ebenfalls als Kontrollvariable (Datenquelle [ESTV \(2024\)](#)).

Durchschnittliche Wohnfläche Haushalte: Diese Variable kombiniert die Bevölkerungsstatistiken aus der STATPOP-Erhebung des Bundesamts für Statistik (2020) mit den Zahlen zu den Wohnungsflächen aus der Gebäude- und Wohnungsstatistik des BFS (2024). Beide Datensätze sind auf Hektarebene ($100\text{m} \times 100\text{m}$) verfügbar. Es wird jeweils pro Hektar die totale Wohnfläche dividiert durch die Anzahl Haushalte. Für eine angemessene Skalierung und eine vereinfachte Interpretation der Ergebnisse wird die resultierende Zahl durch 10 geteilt. Dadurch werden die grundlegenden Beziehungen nicht verändert und somit werden auch die Regressionsergebnisse (vgl. Tabelle 2) nicht verzerrt.

Anteil Ausländer (Nicht-EU): Eine weitere Kontrollvariable stellt der Anteil an Personen mit aussereuropäischen Staatsangehörigkeiten im Hektar der Wohnadresse gemäss STATPOP-Erhebung dar (BFS, 2020). Die spezifische Einkommensanalyse zeigt, dass ein höherer Anteil an Nicht-EU Ausländern mit einem geringeren Anteil an hohen Einkommen einhergeht.

Anteil Einfamilienhäuser: Ferner wird für den Anteil der Einfamilienhäuser an der totalen Anzahl Häuser im Hektar der Wohnadresse gemäss Gebäude- und Wohnungsstatistik kontrolliert (BFS, 2024).

Im Modellierungsprozess wurden weitere Variablen überprüft. Diese wiesen jedoch keine Signifikanz auf und wurden zum Teil ersetzt. Diese sind u.a.:

² Die Berufsinformationen stammen aus alten Dateneinträgen der Individuen vor der Pensionierung.

³ Das negative Vorzeichen bei den Altersgruppen im Regressionsoutput in Tabelle 2 ergibt sich aufgrund der unterschiedlichen Gegebenheiten in den anderen berücksichtigten Variablen, beispielsweise dem Anteil der IPV-Bezüger, nicht aber aufgrund der Tatsache, dass in gewissen Fällen stellvertretend das Familieneinkommen verwendet wird.

Höhe der IPV-Beiträge: Wurde durch den IPV-Dummy ersetzt.

Anteil Schweizer / Ausländer je Hektar: Wurde durch den Anteil Nicht-EU-Ausländer ersetzt.

Durchschnittliche Wohnungsgrösse je Hektar: Wurde durch die durchschnittliche Wohnfläche pro Haushalt je Hektar ersetzt.

VVG-Kennzahlen (Anzahl Produkte, Dummies für einzelne Produktgruppen (z.B. VVG-Spitalversicherung)): Wurde durch die VVG-Prämien ersetzt.

Akademische Titel: Aufgrund von zu wenigen Angaben wurde diese Variable weggelassen.

1.4 Prognosestärke

Um die Qualität des Modells und dessen Klassifizierungen in Hoch- und Niedriglohnberufe besser beurteilen zu können, wurde der Datensatz in einen Trainings- und einen Testdatensatz unterteilt. Der Trainingsdatensatz dient dazu, Muster und intervariable Beziehungen in den Daten zu erkennen. Wie es der Name vermuten lässt, wird also das Modell auf diesem Teil der Daten «trainiert». Der Testdatensatz wird hingegen bei diesen Berechnungen ausgeschlossen und nur für die Beurteilung der Modellperformance herangezogen, damit es zu keiner Überspezifikation des Modells kommt.

Tabelle 1 zeigt das resultierende Mengengerüst für den Trainings- und den Testdatensatz. Das Verhältnis zwischen der Anzahl der Personen mit einem niedrigen Einkommen und der Anzahl der Personen mit einem hohen Einkommen ist bei beiden Datensätzen vergleichbar. Dies stellt eine unverzerrte Evaluation des Modells sicher.

Datensatz	Hohes Einkommen (Anzahl Beobachtungen)	Tiefes Einkommen (Anzahl Beobachtungen)
Trainingsdaten	24'567	61'892
Testdaten	13'431	32'447

Tabelle 1: Mengengerüst Modellierung

In der Abbildung A1 im Anhang ist die ROC-Kurve für den Testdatensatz abgebildet. Die ROC-Kurve ist eine Methode zur Bewertung und Optimierung von Analysestrategien. Sie zeigt unter allen vorgenommenen Klassifizierungen im Testdatensatz das Verhältnis zwischen der Rate der richtigen positiven Ergebnisse (Sensitivität) und der Rate der falschen positiven Ergebnisse für verschiedene Schwellenwerte. Eine ROC-Kurve mit einer grösseren Fläche unter der Kurve weist auf eine bessere Leistungsfähigkeit des Modells hin. Die Fläche unter der Kurve beträgt 0.82. Das heisst, in 82% der Fälle hat eine Person mit einem Hochlohnberuf einen höheren Wert im Einkommensmodell als eine Person mit einem Niedriglohnberuf.

1.5 Resultate Regressionsmodell

Die Tabelle 2 zeigt die Resultate des oben beschriebenen Regressionsmodells. Die Regressionskoeffizienten zeigen, wie die unabhängigen Variablen die Wahrscheinlichkeit einer Hochlohnklassifizierung beeinflussen. Mithilfe des Regressionsmodells wird für alle Individuen ein Wahrscheinlichkeitswert eines Hochlohnberufes berechnet. Anhand dieser Werte wurde eine Einkommensverteilung erstellt, in dem die Individuen nach ihrem Wahrscheinlichkeitswert rangiert werden.

	Regressionskoeffizient (Std. Fehler)
<i>Konstante</i>	-0.86 (0.11) ***
<i>Altersgruppe</i>	
21 - 25	-0.67 (0.08) ***
26 - 30	-0.61 (0.08) ***
31 - 35	-0.32 (0.07) ***
36 - 40	-0.12 (0.07) .
41 - 45	-0.02 (0.07)
46 - 50	0.01 (0.07)
51 - 55	-0.23 (0.08) **
56 - 60	-0.24 (0.08) **
61 - 65	-0.29 (0.08) ***
66+	-0.02 (0.08)
<i>Franchise</i>	
CHF 500.-	-0.15 (0.04) ***
CHF 1000.-	0.17 (0.06) **
CHF 1500.-	0.51 (0.03) ***
CHF 2000.-	0.84 (0.05) ***
CHF 2500.-	1.01 (0.02) ***
<i>Andere</i>	
Einkommensindex	2.63 (0.04) ***
IPV-Empfänger	-1.29 (0.03) ***
Monatsprämie VVG	0.00 (0.00) ***
Median-Einkommen Gemeinde	0.07 (0.01) ***
Durchschn. Wohnfläche Haushalte	0.03 (0.00) ***
Anteil Ausländer (Nicht-EU)	-1.51 (0.10) ***
Anteil Einfamilienhäuser	0.38 (0.03) ***

Signifikanzniveaus 0'***' 0.001'***' 0.01'**' 0.05'.'

Tabelle 2: Regressionsoutput

1.6 Definition Einkommensgruppen und deskriptive Statistiken

Nachdem die Wahrscheinlichkeitswerte für alle Individuen berechnet und die Beobachtungen danach rangiert sind, lassen sich die Einkommensgruppen bilden, für welche die Umverteilungswirkungen bestimmt werden sollen. Die untersten 25% in der Einkommensverteilung werden den untersten Einkommen zugeordnet. Das ist der Viertel der Beobachtungen mit den tiefsten Hochlohn-Wahrscheinlichkeiten. Individuen zwischen dem 1. Quartil und dem Median gehören zum unteren Mittelstand, diejenigen zwischen dem Median und dem 3. Quartil werden als oberer Mittelstand bezeichnet. Die Datenpunkte mit den 25% höchsten Wahrscheinlichkeitswerten werden als oberste Einkommen definiert. Die Tabellen 3 und 4 zeigen deskriptive Statistiken zu den vier Einkommensgruppen.

Einkommensgruppe	Ø -Einkommen	Ø -Alter	Ø -Bruttoleistungen	Ø -Prämienhöhe
Unterste Einkommen (0-25%)	12381	47	5806	4622
Unterer Mittelstand (25-50%)	42252	50	4662	4325
Oberer Mittelstand (50-75%)	70161	50	3692	4087
Oberste Einkommen (75-100%)	158936	53	4112	4262

Tabelle 3: Deskriptive Statistiken

Einkommensgruppe	Anteil Rentner	Anteil Jugendliche (19-25)	Anteil Personen mit mind. einer PCG*	Anteil Spitalflag**	Anteil Personen mit freier Arztwahl	Anteil Personen mit Franchise 300.-
Unterste Einkommen (0-25%)	20,8%	15,1%	23,9%	8,6%	27,3%	67,1%
Unterer Mittelstand (25-50%)	24,3%	9,6%	21,2%	6,7%	24,0%	50,7%
Oberer Mittelstand (50-75%)	22,9%	7,1%	16,5%	5,0%	21,4%	34,8%
Oberste Einkommen (75-100%)	29,2%	4,4%	16,7%	5,4%	28,9%	26,7%

Tabelle 4: Deskriptive Statistiken

* Pharmazeutische Kostengruppen (PCG) als Morbiditätsindikator zur Ermittlung des Anteils chronisch kranker Versicherter.

** Kategorisierung der Spitalaufenthalte gemäss Definition im Risikoausgleich zur Ermittlung des Anteils Versicherter mit Folgekosten.

Nachdem jedes Individuum in eine Einkommensgruppe eingeteilt worden ist, lassen sich die Umverteilungsströme zwischen diesen Einkommensgruppen ermitteln (vgl. nächster Abschnitt).

2 Umverteilungskomponenten

Die SID Health zeigt die Umverteilung im gesamten OKP-System zwischen den Einkommensgruppen «Unterste Einkommen», «Unterer Mittelstand», «Oberer Mittelstand» und «Oberste Einkommen» (vgl. Abschnitt 1.6). Diese ergibt sich aus dem Bezug aller OKP-Leistungen und der Finanzierung via Steuern und Prämienzahlungen.

Die Umverteilung wird jeweils zwischen den einzelnen Einkommensgruppen mit der gleichen Basisformel berechnet. Die Umverteilung pro Kanton k und Jahr t pro Einkommensgruppe j ist die Abweichung des Durchschnittswerts $x_{j,k,t}$ der Einkommensgruppe zum Durchschnittswert über alle Einkommensklassen $x_{k,t}$.

$$U_{j,k,t} = \frac{x_{j,k,t}}{n_{j,k,t}} - \frac{x_{k,t}}{n_{k,t}}$$

Für die SID Health wird die Umverteilung durch die Prämienzahlungen (Frage 2 der SID Health), die individuellen Prämienverbilligungen (Frage 3 der SID Health) und der kantonalen Finanzierung der stationären Spitalleistungen (Frage 4 der SID Health) berechnet. Die beiden letzten Komponenten werden zusammen mit dem «Versicherungstechnischen Ergebnis» addiert (vgl. Abschnitt 2.1 weiter unten), um die gesamte Umverteilung in der OKP (Frage 1 der SID Health) zu berechnen.

2.1 Gesamte OKP-Umverteilung

Um die gesamte Umverteilung in der OKP zu berechnen, wird die Umverteilung durch die einzelnen Komponenten (Versicherungstechnisches Ergebnis, individuelle Prämienverbilligungen und kantonale Finanzierung der Spitalbeiträge) addiert. Wiederum wird der durchschnittliche Wert des Totals a pro Einkommensgruppe j , Kanton k und Jahr t mit dem Durchschnitts-Total aller Gruppen in einem bestimmten Kanton k und Jahr t verglichen.

$$U_{j,k,t} = \frac{a_{j,k,t}}{n_{j,k,t}} - \frac{a_{k,t}}{n_{k,t}}$$

Versicherungstechnisches Ergebnis

Das versicherungstechnische Ergebnis ist die OKP-Bilanz eines Individuums. Es wird für Individuum i im Jahr t mittels folgender Formel berechnet:

$$e_{i,t} = p_{i,t} + r_{i,t} - l_{i,t} + k_{i,t},$$

wobei

$p_{i,t}$ die geforderte Nettoprämie, d.h. Prämie abzüglich allfälliger Rabatte für die Wahlfranchise oder Einschränkung der Arztwahl darstellt.

$r_{i,t}$ dem Risikoausgleichsbeitrag entspricht. Dieser ist abhängig von Demografie- und Morbiditätsmerkmalen der versicherten Person. Bei jungen Personen ohne chronische Erkrankungen oder Spitalaufenthalten im Vorjahr resultiert ein negativer Wert (d.h. der Risikoausgleich

ist erlösmindernd). Bei älteren Personen oder Personen mit chronischen Erkrankungen oder Spitalaufenthalten im Vorjahr ist der Betrag positiv.

$l_{i,t}$ den bezogenen Bruttoleistungen, d.h. Leistungen vor Abzug der Kostenbeteiligung, entspricht. Nicht inbegriffen sind die vom Kanton bezahlten Beiträge an die stationären Spitalleistungen.

$k_{i,t}$ die Kostenbeteiligung (Franchise, Selbstbehalt und Spitalbeitrag) an den bezogenen Leistungen des Individuums ist.

Für die Berechnung des versicherungstechnischen Ergebnisses der Einkommensklassen e_j können die Werte der Individuen aus den einzelnen Einkommensklassen aufsummiert werden. Anschliessend wird die Umverteilung $U_{j,k,t}$ mit der generellen Formel berechnet:

$$U_{j,k,t} = \frac{e_{j,k,t}}{n_{j,k,t}} - \frac{e_{k,t}}{n_{k,t}}$$

Zu der Umverteilung durch das versicherungstechnische Ergebnis wird die Umverteilung durch die individuellen Prämienverbilligungen und die kantonale Finanzierung an den stationären Spitalleistungen addiert. Daraus resultiert die gesamthafte OKP-Umverteilung, die den Bezug aller KVG-pflichtigen Leistungen und deren Finanzierung durch Prämien und Steuern berücksichtigt (Frage 1 der SID Health).

2.2 Prämienzahlungen

OKP-Versicherte bezahlen je nach gewählter Franchisenhöhe, Alter und Kanton unterschiedlich hohe Prämien, wodurch es zur Umverteilung zwischen den Einkommensgruppen kommt. Die SID Health zeigt dies anhand der geforderten Nettoprämien p , die via das versicherungstechnische Ergebnis in die OKP-Gesamtumverteilung einfließen. Für die Berechnung der resultierenden Umverteilung zwischen den Einkommensgruppen werden jeweils die durchschnittlichen Prämienzahlungen p pro Einkommensgruppe j und Kanton k im Jahr t vom Durchschnitt der Prämienzahlungen p aller Prämienzahlungen im Kanton k und Jahr t abgezogen.

$$U_{j,k,t} = \frac{p_{j,k,t}}{n_{j,k,t}} - \frac{p_{k,t}}{n_{k,t}}$$

2.3 Individuelle Prämienverbilligungen (IPV)

Individuen können selber individuelle Prämienverbilligungen (IPV) erhalten und finanzieren diese indirekt via Steuern. Die Umverteilung, die durch den Bezug oder respektive durch die Finanzierung der IPV resultiert, wird anhand des Nettowerts pro Gruppe ermittelt. Bezieht eine Gruppe durchschnittlich mehr IPV als sie via Steuern zur Finanzierung beiträgt, ist der Nettowert negativ. Das Umgekehrte gilt, wenn mehr finanziert als bezogen wird. Die Umverteilung wird folglich bestimmt, in dem die Abweichung zwischen dem durchschnittlichen IPV-Nettowert i einer Einkommensgruppe j im Kanton k und Jahr t und dem Durchschnitt des IPV-Nettowerts i über alle Einkommensgruppen im Kanton k und Jahr t berechnet wird.

$$U_{j,k,t} = \frac{i_{j,k,t}}{n_{j,k,t}} - \frac{i_{k,t}}{n_{k,t}}$$

Die Berechnung der Finanzierungsseite des Netto-IPV-Wertes erfolgte separat mit Daten der Eidgenössischen Steuerverwaltung und wird im Kapitel 3 des vorliegenden technischen Appendix erklärt.

2.4 Kantonale Finanzierung der stationären Spitalleistungen

Die Kantone bezahlen einen fixen Anteil an den stationären Leistungen von OKP-Versicherten, also von jenen Gesundheitsleistungen, die bei einem stationären Spitalbesuch anfallen. Dies führt dazu, dass Individuen durch einen stationären Spitalaufenthalt Leistungen beziehen, die via Kantonssteuern finanziert werden. Analog zur Umverteilung durch die IPV wird der Nettowert zwischen Finanzierung und Bezug verwendet. So ergibt sich die Umverteilung durch die Differenz zwischen dem durchschnittlichen Netto-Spitalbeitrag s pro Einkommensgruppe j in ihrem Wohnkanton k und Jahr t zum Durchschnitt über alle Einkommensgruppen in demselben Kanton k und Jahr t .

$$U_{j,k,t} = \frac{s_{j,k,t}}{n_{j,k,t}} - \frac{s_{k,t}}{n_{k,t}}$$

Die Berechnung der Finanzierungsseite der kantonalen Spitalbeiträgen mit Daten der Eidgenössischen Steuerstatistik (ESTV) wird in Kapitel 3 genauer beschrieben.

3 Berechnung Steuerfinanzierung

Die Umverteilung bei den individuellen Prämienverbilligungen (IPV) und kantonalen Spitalbeiträgen entsteht nicht nur durch den unterschiedlichen Leistungsbezug der Einkommensgruppen. Die Einkommensgruppen tragen auch unterschiedlich zur Finanzierung mittels Steuern bei. Zur Berechnung der Finanzierungseite werden daher folgende Statistiken der Eidgenössischen Steuerverwaltung verwendet:

- Die Tabelle 2b aus den [Statistiken zur direkten Bundessteuer](#) liefert Informationen zur Steuerbelastung auf Bundesebene für 21 Einkommensklassen, inklusive Mengengerüst pro Kanton.
- Die [Steuerbelastungsstatistik](#) liefert Informationen zur Steuerbelastung auf Kantons-ebene nach Subjekt und Einkommen.

Aus den Statistiken zur direkten Bundessteuer können die Perzentile der einzelnen Einkommensklassen sowie das Durchschnittseinkommen und der durchschnittliche Steuerertrag pro Einkommensgruppe abgeleitet werden. Ebenfalls lässt sich der Anteil an verheirateten Personen pro Einkommensklasse ermitteln. So erhält man für eine bestimmte Einkommensklasse j im Kanton k folgende Kennzahlen:

- $(p_{jk}^{\min}, p_{jk}^{\max})$: Das Perzentilintervall der Einkommensklasse
- \bar{i}_{jk} : Das Durchschnittseinkommen der Einkommensklasse
- \overline{se}_{jk} : Der durchschnittliche Bundessteuerertrag der Einkommensklasse
- s_{jk}^{verh} : Der Anteil verheiratete Personen in der Einkommensklasse

Anschliessend werden die Steuerbeträge der Bundes- und Kantonssteuern aus der Steuerbelastungsstatistik mit Hilfe des Durchschnittseinkommens \bar{i}_{jk} den einzelnen Einkommensklassen zugewiesen. In der Steuerbelastungsstatistik sind nur die Steuerbeträge für ausgewählte Einkommen $\{i_0, \dots, i_n\}$ enthalten. Der Einkommensklasse j im Kanton k wird dann der Steuerbetrag für das Einkommen i_{jk} zugewiesen, wenn $\frac{i_{j-1k} + i_{jk}}{2} \leq \bar{i}_{jk} < \frac{i_{jk} + i_{j+1k}}{2}$ gilt.

Anhand dieser Hochrechnungen lassen sich folgende Grössen pro Einkommensklasse j im Kanton k ermitteln:

- $\text{kanton}_{jk}^{\text{verh}}$: Der Kantonssteuerbetrag für verheiratete Personen (ohne Doppelinkommen und ohne Kinder)
- $\text{kanton}_{jk}^{\text{ledig}}$: Der Kantonssteuerbetrag für ledige Personen (ohne Kinder)
- $\text{bund}_j^{\text{verh}}$: Der Bundessteuerbetrag für verheiratete Personen (ohne Doppelinkommen und ohne Kinder)
- $\text{bund}_j^{\text{ledig}}$: Der Bundessteuerbetrag für ledige Personen (ohne Kinder)

Anschliessend wird ein gewichtetes Mittel aus den Steuerbeträgen für Verheiratete und Unverheiratete berechnet, wobei der Anteil an verheirateten Personen in den 21 Einkommensklassen jeweils als Gewicht dient:

- $\text{kanton}_{jk} = s_{jk}^{\text{verh}} \times \text{kanton}_{jk}^{\text{verh}} + (1 - s_{jk}^{\text{verh}}) \times \text{kanton}_{jk}^{\text{ledig}}$
- $\text{bund}_{jk} = s_{jk}^{\text{verh}} \times \text{bund}_{jk}^{\text{verh}} + (1 - s_{jk}^{\text{verh}}) \times \text{bund}_{jk}^{\text{ledig}}$

Zum Schluss ist die Berechnung zweier Faktoren nötig:

- IPV: Die Prämienverbilligung wird teilweise vom Bund finanziert. Deshalb wird – gewichtet nach den IPV-Anteilen des Kantons – ein Durchschnitt aus Bundes- und Kantonssteuerbelastung berechnet.

$$Faktor_{IPV} = \frac{Anteil_{IPV} * kanton_{jk} + (1 - Anteil_{IPV}) * bund_{jk}}{bund_{jk}}$$

- Kantonsbeitrag an den stationären Spitalleistungen: Da die Kantonsbeiträge von den Kantonen finanziert werden und ein Teil der Kantonsfinanzen aus Bundestransfers besteht, muss um diesen Faktor korrigiert werden. Dazu wird das Verhältnis zwischen Kantons- und Bundessteuern bestimmt.

$$Faktor_{Kanton} = \frac{kanton_{jk}}{bund_{jk}}$$

Die durchschnittlichen Steuererträge aus der Bundessteuer \overline{se}_{jk} werden anschliessend mit diesen Faktoren multipliziert, um eine durchschnittliche Beitragseinheit \overline{i}_{jk} zur IPV resp. dem Kantonsbeitrag zu den Spitalbeiträgen \overline{sb}_{jk} zu berechnen.

$$\overline{i}_{jk} = Faktor_{IPV} * \overline{se}_{jk}$$

$$\overline{sb}_{jk} = Faktor_{Kanton} * \overline{se}_{jk}$$

Die Individuen aus dem CSS-Kollektiv werden basierend auf den Perzentilwerten des oben beschriebenen Einkommensmodells in Einkommensklassen eingeteilt (über die Perzentilintervalle der Einkommensklassen aus der Steuerstatistik). Die entsprechenden Pro-Kopf-Steuerbeträge werden dadurch den Individualbeobachtungen zugewiesen. Daraus lassen sich schliesslich die Anteile an der Finanzierung der IPV und an den Kantonsbeiträgen der vier Einkommensgruppen bestimmen, indem die Beitragseinheiten innerhalb der vier Gruppen aufsummiert werden.

4 Literaturverzeichnis

BFS, Bundesamt für Statistik (2020). Statistik der Bevölkerung und Haushalte (STATPOP), Geodaten 2019. Online verfügbar: <https://www.bfs.admin.ch/asset/de/14716365>. Zugriff am 04.09.2024.

BFS, Bundesamt für Statistik (2024). Gebäude- und Wohnungsstatistik (seit 2009). Online verfügbar: <https://www.bfs.admin.ch/bfs/de/home/statistiken/bau-wohnungswesen/erhebungen/qws2009.html>. Zugriff am 04.09.2024.

ESTV, Eidgenössische Steuerverwaltung (2024). Statistiken zur direkten Bundessteuer (Dbst). Online verfügbar: <https://www.estv.admin.ch/estv/de/home/die-estv/steuerstatistiken-estv/allgemeine-steuerstatistiken/direkte-bundessteuer.html>. Zugriff am 04.09.2024.

Schmid, C. P., Schreiner, N., & Stutzer, A. (2022). Transfer Payment Systems and Financial Distress: Insights from Health Insurance Premium Subsidies. *Journal of the European Economic Association*, 20(5), 1829-1858.

Anhang

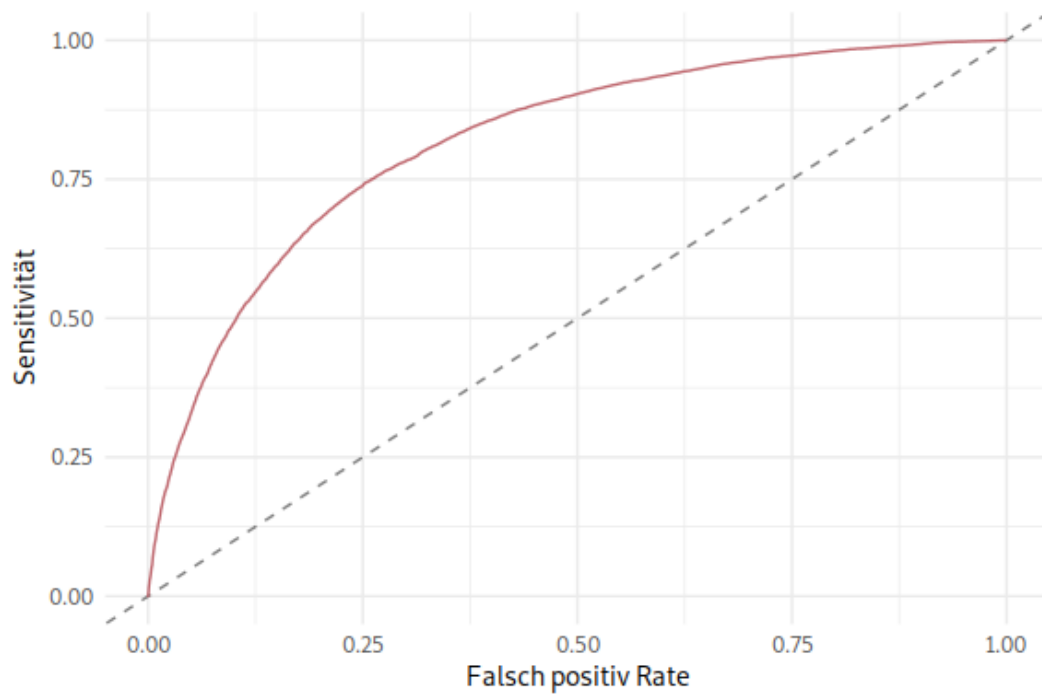


Abbildung A1: ROC-Kurve für den Testdatensatz