

Taxpayer-Panel 2 (TPP2)

– Lohn- und Einkommensteuerstatistik 2012-2020 –

Moritz Wittmaack

1 Einführung

Seit seiner ersten Bereitstellung im Jahr 2007 bietet das Taxpayer-Panel (TPP) Forschenden die Möglichkeit, die Daten der Lohn- und Einkommensteuerstatistik (LESt) im Längsschnitt zu analysieren.¹ Hierzu wurden Beobachtungen derselben Steuerpflichtigen aus den Querschnittsdaten über die Zeit verknüpft. Von 2001 bis 2011 wurden die Daten der jährlichen Einkommensteuerstatistik (Geschäftsstatistik) und ab 2012 die Daten der jährlichen Lohn- und Einkommensteuerstatistik (Bundesstatistik) genutzt. Der Wissenschaft wurden diese Verknüpfungen als unterschiedliche Forschungsdatenbestände – projektspezifischer Zuschnitt des TPP-Vollmaterials, 5%-Stichprobe und 0,5%-Stichprobe – jeweils in Form eines breiten Datensatzes über die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (FDZ) zur Verfügung gestellt. In jeder Zeile des projektspezifischen Vollmaterialzuschnitts wurden dabei die Beobachtungen eines Steuerpflichtigen erfasst, wenn dieser im Beobachtungszeitraum mindestens zweimal in den jährlichen Querschnitten vorhanden war und verknüpft werden konnte. Die im Beobachtungszeitraum mindestens fünfmal verknüpfbaren Steuerpflichtigen bildeten die Zeilen, auf deren Grundlage die geschichtete 5%-Stichprobe gezogen wurde. Diese Zeilen bildeten wiederum die Grundlage der einfachen Zufallsstichprobe für den 0,5%-Stichproben-Forschungsdatenbestand.²

Die Selektivität der jeweiligen TPP-Forschungsdatenbestände schränkte die Analysemöglichkeiten jedoch teilweise ein, etwa bei Nutzung der Stichproben. Wenn beispielsweise die Entwicklung von Steuerpflichtigen nach Wechsel von einer Zusammenveranlagung in eine Einzelveranlagung (oder anders herum) im Fokus des Forschungsinteresses stand, musste häufig auf einen projektspezifischen Zuschnitt des Vollmaterials zurückgegriffen werden, was ursprünglich nur als Ausnahme vorgesehen war.³ Dies führte zu einem erheblichen Mehraufwand auf Seiten des

¹ Siehe beispielsweise Gerber & Kann (2019) und Meyer (2022) für Längsschnittanalysen mit dem TPP. Weitere Publikationen auf Basis amtlicher Mikrodaten finden Sie auch unter: <https://www.forschungsdatenzentrum.de/de/veroeffentlichungen>

² Siehe Kriete-Dodds & Vorgrimler (2007) für eine Beschreibung der Methodik zum Zeitpunkt der erstmaligen Bereitstellung des TPP sowie Gerber & Hammer (2013) für die methodischen Weiterentwicklungen des TPP und Statistisches Bundesamt (2024) für das Nutzungskonzept des TPP über den Beobachtungszeitraum 2001-2020.

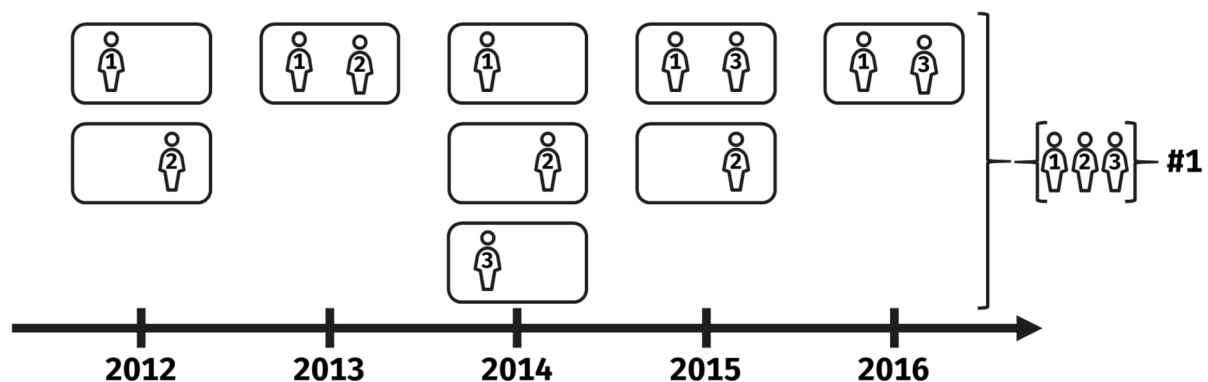
³ Ein Steuerpflichtiger besteht bei einer Einzelveranlagung aus einem Steuerfall ($Stpfl = Stfall_A = A$ oder $Stpfl = Stfall_B = B$) und bei einer Zusammenveranlagung aus zwei Steuerfällen ($Stpfl = [Stfall_A, Stfall_B] = [A, B]$). Wenn ein Wechsel von einer Zusammenveranlagung in zwei Einzelveranlagungen im Beobachtungszeitraum vorlag, konnte nur einer der vormaligen zusammenveranlagten Steuerfälle als Beobachtung in diesem Steuerpflichtigen fortgeführt werden. Der „frei gewordene“ Steuerfall wurde dann zu einem eigenständigen Steuerpflichtigen, wobei dessen Historie als zusammenveranlagter Steuerfall nicht nachvollziehbar war (z. B. $Stpfl_1 = [A, B]_{t_0}, A_{t_1}$ und $Stpfl_2 = B_{t_1}$). Ebenfalls zu einer Zensur führte die Beobachtung eines Wechsels in eine Zusammenveranlagung: Wenn ein Wechsel von zwei Einzelveranlagungen in eine Zusammenveranlagung im Beobachtungszeitraum vorlag, konnte die Zusammenveranlagung nur einem der vormaligen einzelveranlagten Steuerpflichtigen als Beobachtung zugeordnet werden. Der „frei bleibende“ Steuerpflichtige blieb dann ein eigenständiger Steuerpflichtiger, wobei dessen Entwicklung in eine Zusammenveranlagung nicht nachvollziehbar war (z. B. $Stpfl_3 = A_{t_0}, [A, B]_{t_1}$ und $Stpfl_4 = B_{t_0}$). Bei der Nutzung der Stichproben war bisher nicht sichergestellt, dass alle durch eine jemals beobachtete Zusammenveranlagung derselben Steuerfälle in Beziehung zueinanderstehenden Steuerpflichtigen zusammen selektiert wurden. Daneben führte auch die Beschränkung der mindestens zweimaligen Beobachtung desselben Steuerpflichtigen im Beobachtungszeitraum bei der Selektion der das Vollmaterial bildenden Menge an Steuerpflichtigen dazu, dass eine generelle Zensur der Daten bei der Beobachtung des Wechsels von Zusammenveranlagung in eine Einzelveranlagung (oder anders herum) nicht auszuschließen war. In den zuvor genannten Beispielen wären $Stpfl_2$ und $Stpfl_4$ nicht Teil des Vollmaterials geworden.

Statistischen Bundesamts, da jedes Vollmaterial aufgrund der zu verarbeitenden Datenmengen mit einer Beschränkung auf 25 Merkmale pro Statistikjahr projektspezifisch neu erstellt werden musste. Damit gingen zudem Wartezeiten für die Forschenden zwischen der Einreichung des Forschungsantrags und den ersten Analysen einher.

Mit dem Taxpayer-Panel 2 (TPP2), welches erstmals für den Beobachtungszeitraum 2012-2020 bereitgestellt wird, konnten diese designbedingten Unschärfen überwunden werden. An die Stelle der Verknüpfung von Beobachtungen desselben Steuerpflichtigen über die Zeit als zentrale Einheit der Datenaufbereitung und Stichprobenziehung treten im weiteren die als Cluster bezeichneten Zusammenhänge von Steuerpflichtigen im Beobachtungszeitraum. Ein Cluster stellt die Gesamtheit aller Beobachtungen von Steuerpflichtigen dar, welche sich aus sukzessiven Zusammenveranlagungen derselben Steuerfälle im Beobachtungszeitraum ergeben. Wenn im Beobachtungszeitraum kein Veranlagungswechsel der in einem Steuerpflichtigen enthaltenen Steuerfälle beobachtet wird, dann besteht dieses Cluster lediglich aus den vorliegenden Beobachtungen dieses Steuerpflichtigen.

Abbildung 1 stellt ein Cluster schematisch dar: Die Personen 1 und 2 bilden im Jahr 2013 die beiden Steuerfälle einer Zusammenveranlagung. In den Jahren 2015 und 2016 bilden Person 1 und 3 ebenfalls die beiden Steuerfälle einer Zusammenveranlagung. Person 1 bildet in diesem Beispiel somit das verbindende Element der Personen 2 und 3. Folglich bilden alle neun Steuerpflichtigen, welche im Beobachtungszeitraum 2012 bis 2016 auf die Personen 1, 2 und 3 als Steuerfälle zurückgehen, ein Cluster. Sofern dieses Cluster in die Stichprobe gezogen würde, stünden alle diese neun Steuerpflichtigen für Analysen zur Verfügung.

Abbildung 1: Schematische Darstellung eines Clusters



Das TPP2 wird der Forschung in Form zweier Datenbestände – 10%-Stichprobe und 1%-Stichprobe – zur Verfügung gestellt, wobei der bisherige zweijährige Veröffentlichungsturnus beibehalten wird. Im TPP enthielt jede Zeile die Informationen eines Steuerpflichtigen über den gesamten Beobachtungszeitraum, wobei die Informationen der einzelnen Statistikjahre nebeneinander aufgeführt waren (breiter Datensatz). Im TPP2 enthält nun jede Zeile die Beobachtung eines Steuerpflichtigen eines Statistikjahres, sofern eine Beobachtung des Steuerpflichtigen für das Statistikjahr vorliegt (langer Datensatz). Durch Vergabe eines persistenten Identifikators für jeden Steuerfall können sich die Forschenden selbstständig eine auf das Forschungsinteresse angepasste Datensatzstruktur erstellen. Da erstens keine Beschränkung bezüglich der Mindestanzahl von Beobachtungen zur Selektion der Steuerpflichtigen in das zur Stichprobenziehung genutzte TPP2-Vollmaterial Anwendung findet, zweitens die geschichtete Stichprobe aus den Clustern gezogen wird und drittens der Stichprobenumfang deutlich erhöht wurde, ist ein Rückgriff auf das TPP2-Vollmaterial für die eingangs erwähnten Analysen nicht mehr notwendig. Hierdurch kann zum einen die Zeitspanne zwischen Beantragung der Forschungsdaten

und dem Start der Datenanalyse verkürzt werden. Zum anderen erhöht sich das Analysepotential insgesamt gegenüber dem bisherigen TPP, welches letztmalig für den Beobachtungszeitraum 2001-2020 zur Verfügung gestellt wird.

2 Datenmaterial

Die seit 2012 bestehende jährliche LEST bildet die Datengrundlage des TPP2. Jeder Querschnitt der jährlichen Bundesstatistik zur LEST setzt sich aus ca. 29 Millionen veranlagten Steuerpflichtigen und ca. 13 Millionen nicht veranlagten Steuerpflichtigen, den sogenannten Lohnsteuerfällen, zusammen.⁴

Der Beginn des Beobachtungszeitraums für das TPP2 wurde aus mehreren Gründen auf 2012 gesetzt: Erstens stehen die eindeutigen und persistenten Steueridentifikationsnummern (IdNr) zur Bestimmung der Cluster durchgehend seit 2012 zur Verfügung. Zweitens werden die Daten der Bundesstatistik zur LEST für alle Bundesländer einheitlich im Steuerstatistischen Gesamtsystem (StSys) aufbereitet.⁵ Damit sind auch für den gesamten Zeitraum die Lohnsteuerfälle enthalten. Drittens stellt sich nicht das Problem der teilweisen Verschiebungen bei der Belegung von Kennzahlen durch den Wechsel der Datengrundlage (Bundes- anstelle Geschäftsstatistik), welche die Arbeit mit dem TPP teilweise erschwert hatten.

Die Arbeitsschritte zur Erstellung des TPP2 gliedern sich wie folgt:

1. Die Querschnittsdaten der Bundesstatistik zur LEST werden gepoolt.
2. Alle in zusammenveranlagten Steuerpflichtigen enthaltenen distinkten Kombinationen von IdNr werden herausgeschrieben. Diese IdNr-Kombinationen werden dann auf sich überschneidende IdNr geprüft, sodass ggf. aus einer zweigliedrigen IdNr-Kombination eine mehrgliedrige IdNr-Kombination wird (z. B. 1—2 und 2—3 → 1—2—3).
3. Alle ausschließlich einzeln vorkommenden IdNr erhalten eine eigene einzigartige Clusternummer. Alle in einer IdNr-Kombination enthaltenen IdNr erhalten eine gemeinsame einzigartige Clusternummer. Die Clusternummern werden dann über die IdNr der Steuerfälle auf die Steuerpflichtigen im gepoolten Querschnitt übertragen.
4. Die Geschlechts- und die Geburtsjahresinformationen werden für jede IdNr harmonisiert und gegebenenfalls imputiert. Auch werden fehlende oder fehlerhafte IdNr innerhalb der Cluster imputiert.
5. Als Vorbereitung der Stichprobenziehung werden die Steuerpflichtigen eines Jahres in eine von 13 Schichten pro Bundesland eingeteilt. So entstehen die insgesamt $16 * 13 = 208$ Schichten, welche in Tabelle 1 abgetragen sind.

⁴ Siehe Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (2022) für allgemeine und methodische Informationen zur Lohn- und Einkommensteuerstatistik.

⁵ Siehe Schneider (2012a, 2012b) für Erläuterungen zum StSys.

Tabelle 1: Schichtplan des TPP2

Bundesland	Veranlagt / Nicht veranlagt	Grund- / Splittingtabelle	Mann / Frau / Paar	Überwiegende Einkunftsart	Schicht
Schleswig- Holstein	Veranlagt	Grundtabelle	Mann	Gewinn-EK	1
				Kapital-EK oder VuV	2
				Nslbst. Arbeit oder Sonstige EK	3
			Frau	Gewinn-EK	4
				Kapital-EK oder VuV	5
				Nslbst. Arbeit oder Sonstige EK	6
		Splittingtabelle	Paar	Gewinn-EK	7
				Kapital-EK oder VuV	8
				Nslbst. Arbeit oder Sonstige EK	9
	Nicht veranlagt	Grundtabelle	Mann	Nslbst. Arbeit	10
			Frau	Nslbst. Arbeit	11
		Splittingtabelle	Mann	Nslbst. Arbeit	12
			Frau	Nslbst. Arbeit	13
...
Thüringen	Nicht veranlagt	Splittingtabelle	Frau	Nslbst. Arbeit	208

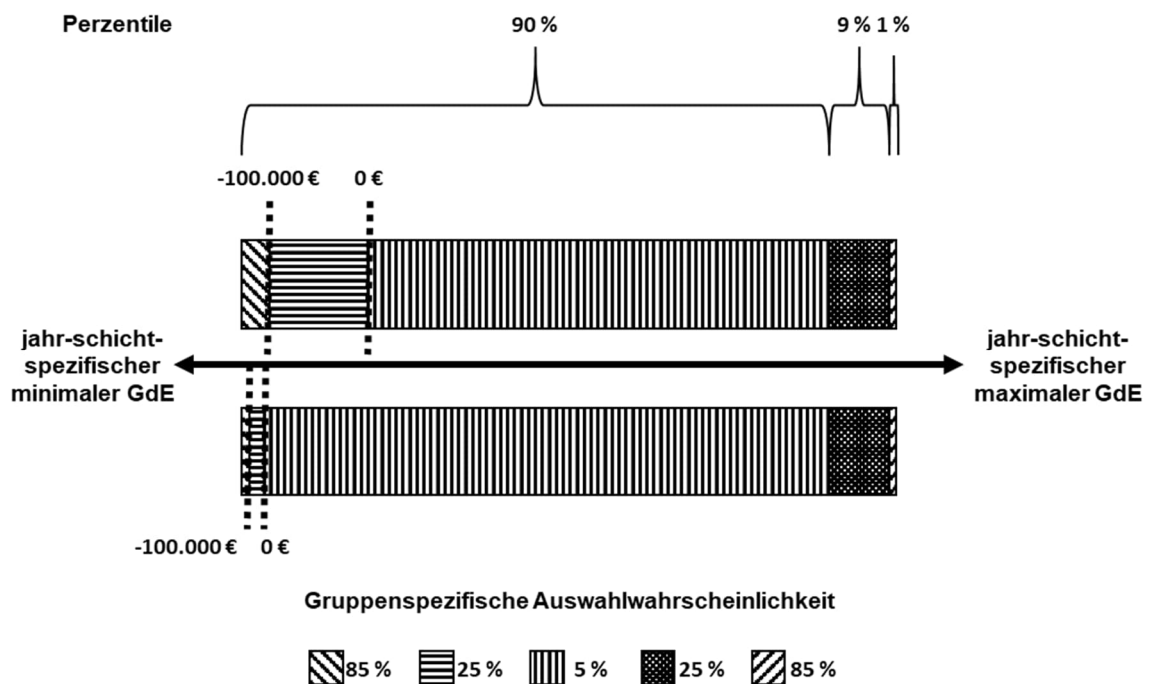
(Anmerkung: Die Abkürzungen sind folgendermaßen zu verstehen: „Gewinn-EK“: Gewinneinkünfte, d. h. Einkünfte aus Land- und Forstwirtschaft, Einkünfte aus Gewerbebetrieb und Einkünfte aus selbständiger Arbeit; „Kapital-EK oder VuV“: Einkünfte aus Kapitalvermögen oder aus Vermietung und Verpachtung; „Nslbst. Arbeit oder Sonstige EK“: Einkünfte aus nichtselbständiger Arbeit und sonstige Einkünfte; „Nslbst. Arbeit“: Einkünften aus nichtselbständiger Arbeit.)

Hiernach werden die Steuerpflichtigen jeder Jahr-Schicht-Kombination nach dem Gesamtbetrag der Einkünfte (GdE) aufsteigend sortiert. In Abhängigkeit von der Höhe des GdE werden die Steuerpflichtigen in fünf disjunkte Gruppen eingeteilt und ihnen wird eine gruppenspezifische Auswahlwahrscheinlichkeit zugeteilt:

- Alle Steuerpflichtigen mit einem GdE von mindestens 0 € bis zum 90. Perzentil (Gruppe 3) erhalten eine Auswahlwahrscheinlichkeit von 5 %.
- Alle Steuerpflichtigen mit einem negativen GdE bis einschließlich - 100.000 € (Gruppe 2) sowie alle Steuerpflichtigen zwischen dem 90. Perzentil und dem 99. Perzentil (Gruppe 4) erhalten eine Auswahlwahrscheinlichkeit von 25 %.
- Die Steuerpflichtigen mit einem negativen GdE kleiner als - 100.000 € (Gruppe 1) sowie alle Steuerpflichtigen des 99. Perzentils (Gruppe 5) erhalten eine Auswahlwahrscheinlichkeit von 85 %.

Abbildung 2 stellt die Zuordnung der Auswahlwahrscheinlichkeit schematisch für die Steuerpflichtigen von zwei fiktiven Jahr-Schicht-Kombinationen dar. Durch die höheren Auswahlwahrscheinlichkeiten an den Rändern der Verteilung soll sichergestellt werden, dass diese relativ seltenen Beobachtungen in den Forschungsdaten überrepräsentiert sind und somit in ausreichend hoher Fallzahl für Analysen zur Verfügung stehen.

Abbildung 2: Schematische Darstellung zur Vergabe der Auswahlwahrscheinlichkeit im TPP2



Nach der Zuteilung der Auswahlwahrscheinlichkeiten wird innerhalb jedes Clusters der Modalwert der Schichtzuordnung(en) sowie der Modalwert der Gruppe(n) mit der höchsten Auswahlwahrscheinlichkeit bestimmt. Wenn es mehrere Modalwerte gibt, wird zufällig einer von diesen ausgewählt. Aus der Kombination dieser Modalwerte ergibt sich für jedes Cluster eine der $208 * 5 = 1.040$ endgültigen Schichten, aus welcher die geschichtete Stichprobe gezogen wird. Tabelle 2 demonstriert dies beispielhaft: Da zwei der vier Steuerpflichtigen des Clusters 1 der Schicht 20 zugeordnet wurden (Zeile 3 und 4) und ein Steuerpflichtiger dieses Clusters aufgrund seines GdE innerhalb der Schicht 17 im Jahr 2012 in die Gruppe 5 mit einer Auswahlwahrscheinlichkeit von 85% eingeteilt wurde (Zeile 2), ergibt sich für Cluster 1 die gültige Modalwerte-Kombination 20-5. Da Cluster 2 nur einen Steuerpflichtigen enthält (Zeile 5), ist die Modalwerte-Kombination dieses Clusters 112-3.

Tabelle 2: Beispiel für Bestimmung der Modalwerte-Kombination eines Clusters

#	Jahr	Steuerpflichtiger	Cluster	Schicht	Gruppe	Gruppenspez. AWW	Modalwert: Schicht	Modalwert: Gruppe mit höchster gruppenspez. AWW
1	2012	1	1	1	3	5 %	20	5
2	2012	2	1	17	5	85 %		
3	2013	3	1	20	4	25 %		
4	2014	4	1	20	4	25 %		
5	2012	5	2	112	3	5 %	112	3
6

(Anmerkung: „Gruppenspez. AWW“: Gruppenspezifische Auswahlwahrscheinlichkeit.)

Der Umfang der zu ziehenden Cluster pro Modalwerte-Kombination ergibt sich aus dem Modalwert der Gruppe mit der höchsten Auswahlwahrscheinlichkeit. Um einen gleichbleibenden Stichprobenumfang zu gewährleisten (10 % für die kontrollierte Datenfernverarbeitung [KDFV] im Forschungsdatenzentrum des Bundes und 1 % für die Nutzung an den Gastwissenschaftsarbeitsplätzen [GWAP] in den FDZ des Bundes und der Länder; siehe Tabelle 3), werden solange Cluster ausgewählt, bis in der Stichprobe annähernd 10 % aller Cluster insgesamt enthalten sind und sich die gezogenen Cluster annähernd in dem Umfang auf die Modalwerte-Kombinationen verteilen, der sich aus der modalwertekombinationsspezifischen Auswahlwahrscheinlichkeit ergibt.⁶ Für die Unterstichprobe zur Nutzung an den Gastwissenschaftsarbeitsplätzen in den FDZ werden innerhalb aller Modalwerte-Kombinationen 10 % der bereits selektierten Cluster ausgewählt. Somit ergibt sich ein Umfang von 1% aller vorhandenen Cluster. Die Grundlage des Datenstrukturfiles bildet wiederum eine 0,1%-Unterstichprobe aus der 1%-Stichprobe, wobei nur Cluster mit einer Auswahlwahrscheinlichkeit von 5 % oder 25 % die Grundlage dieser geschichteten Unterstichprobenziehung bilden. Jeder in die Stichproben selektierte Steuerpflichtige besitzt einen clusterspezifischen Hochrechnungsfaktor, wodurch die Stichproben zur Abbildung der gepoolten Querschnitte genutzt werden können.⁷

In der Einkommensteuerstatistik und somit im TPP2 sind fehlende Werte aus Speicherplatzgründen durch „0“ ersetzt. Da die Finanzverwaltungen nicht einheitlich zwischen errechneten „0“-Werten und fehlenden Werten unterscheiden, entsteht dadurch kein Informationsverlust.

3 Nutzungskonzept

Den Anforderungen der Nutzenden aus der Wissenschaft und den IT-technischen Gegebenheiten entsprechend werden drei unterschiedliche Datenprodukte für die Analyse des TPP2 angeboten. Tabelle 3 fasst die drei Produkte zusammen.

⁶ Sofern beispielsweise in einer Modalwerte-Kombination mit einer Auswahlwahrscheinlichkeit von 85 % lediglich 19 Cluster enthalten sind, werden zunächst $\left\lceil \frac{85}{100} * 19 \right\rceil = 17$ Cluster gezogen. Hierdurch werden dann allerdings mehr Cluster gezogen, als anteilig zu ziehen sind, da $\frac{17}{19} \approx 89,5\% > 85\%$. Aus der Summe der selektierten Cluster im Verhältnis zur Gesamtzahl aller Cluster wird dann die Anzahl der wieder aus der 10%-Stichprobe auszusteuern Cluster berechnet. Die auszusteuern Cluster werden dann anteilig auf die Cluster der Modalwerte-Kombinationen verteilt, wobei jede besetzte Modalwerte-Kombination mit mindestens einem Cluster vertreten ist, Diese Korrektur wird für alle drei sukzessiven Stichprobenziehungen durchgeführt.

⁷ Der jeweilige Hochrechnungsfaktor ergibt sich aus der inversen Auswahlwahrscheinlichkeit des selektierten Clusters. Siehe beispielsweise Gabler & Ganninger (2010) zur Nutzung von Hochrechnungsfaktoren.

Tabelle 3: Übersicht des TPP2-Forschungsdatenangebots für den Beobachtungszeitraum 2012-2020

Produkt	Anzahl der Cluster	Anzahl der Beobachtungen	Anzahl der distinkten Steuerfälle	Anonymisierung	Form der Nutzung
Datenstrukturfile	4.711	36.910	6.494	Vollständige Anonymisierung der Daten, keine inhaltlichen Aussagen möglich	Off-Site
1%-Stichprobe	539.028	4.426.638	756.587	Formale Anonymisierung der Daten, pseudonymer amtlicher Gemeindegemeinschaften für Bayern, absolute Anonymisierung der Ergebnisse	Gastwissenschafts-arbeitsplatz in den FDZ
10%-Stichprobe	5.385.420	44.212.116	7.560.155	Formale Anonymisierung der Daten, absolute Anonymisierung der Ergebnisse	Kontrollierte Datenfernverarbeitung im FDZ (Bund)

Die Datensätze des Datenstrukturfiles werden aus einer 0,1%-Unterstichprobe der 1%-Stichprobe gezogen, wobei Cluster mit einer Auswahlwahrscheinlichkeit von 85% nicht berücksichtigt werden. Die absolute Anonymisierung wird durch eine systematische Vertauschung der Geburtsjahresangaben der Steuerfälle sowie einer zufälligen Vertauschung von Merkmalsausprägungen erreicht, wobei die vorkommenden Merkmalsausprägungen getrennt für positive und negative Merkmalsausprägungen vertauscht werden. Hierdurch bleibt die Charakteristik der Variablen erhalten, der Datensatz kann aber in sich inkonsistent sein und lässt auch bei Nutzung des Hochrechnungsfaktors keine inhaltlichen Aussagen zu. Der Datenstrukturfile wird den Nutzenden vorab zur Syntaxentwicklung bereitgestellt.

Für die Nutzung an einem Gastwissenschaftsarbeitsplatz in den Forschungsdatenzentren (FDZ) der Statistischen Ämter des Bundes und der Länder ist die 1%-Stichprobe vorgesehen. Dabei handelt es sich um eine geschichtete Unterstichprobe der formal anonymisierten 10%-Stichprobe. Da der Amtliche Gemeindeschlüssel für bayerische Gemeinden am Gastwissenschaftsarbeitsplatz lediglich in pseudonymisierter Form bereitgestellt werden kann, sind bei diesen Datensätzen die letzten drei Stellen des Amtlichen Gemeindeschlüssels pseudonymisiert.

Tiefer gehende Analysen sind mit der gewichteten 10%-Stichprobe möglich, die über das FDZ des Statistischen Bundesamtes im Rahmen der kontrollierten Datenfernverarbeitung ausgewertet werden kann. Für dieses Produkt wird die vorherige Arbeit mit der 1%-Stichprobe am Gastwissenschaftsarbeitsplatz empfohlen.

Bei Nutzung der kontrollierten Datenfernverarbeitung entstehen für die Nutzung des ersten Gastwissenschaftsarbeitsplatzes keine zusätzlichen Kosten. Die Ergebnisfreigabe erfolgt dabei ausschließlich über die kontrollierte Datenfernverarbeitung. Die Nutzung weiterer Gastwissenschaftsarbeitsplätze oder eines Gastwissenschaftsarbeitsplatzes mit Ergebnisfreigabe ist kostenpflichtig.

Vor Herausgabe werden die Ergebnisse zur Sicherstellung der statistischen Geheimhaltung überprüft. Einzeldaten, wie z. B. minimale und maximale Ausprägungen, werden grundsätzlich gelöscht.⁸ Fragen zum TPP und dem TPP2 werden auch im Forum4MICA unter <https://forum.lifbi.de/c/destatis/12> beantwortet.

Wiesbaden, Dezember 2024

⁸ Hinweise zu den Geheimhaltungsverfahren finden Sie auch unter: <http://www.forschungsdatenzentrum.de/de/geheimhaltung>.

Literatur

- Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (2022): Metadatenreport. Teil I: Allgemeine und methodische Informationen zur Lohn- und Einkommensteuerstatistik (EVAS-Nummer: 73111). Version 1. Wiesbaden. Online verfügbar unter: https://www.forschungsdatenzentrum.de/sites/default/files/lest_mdr-1_v1.pdf [zuletzt abgerufen am 21.11.2024]
- Gabler, S.; Ganninger, M. (2010): „Gewichtung“ In: Christof Wolf; Henning Best (Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden: VS Verlag für Sozialwissenschaften. S. 143-164.
- Gerber, U; Hammer, S. (2013): „Das Taxpayer-Panel 2001 bis 2008“, Wirtschaft und Statistik, 4/2013, S. 729-737.
- Gerber, U.; Kann, K. (2019): „Wer spendet wie viel? Untersuchungen zur Spendenbereitschaft und zur Spendenhöhe mit dem Taxpayer-Panel“, Wirtschaft und Statistik, 6/2019, S. 73-86.
- Kriete-Dodds, S.; Vorgrimler, D. (2007): „Das Taxpayer-Panel der jährlichen Einkommensteuerstatistik“, Wirtschaft und Statistik, 1/2007, S. 77-85.
- Meyer, C. (2022): „Geschlechtsspezifisches Altersvorsorgeverhalten – Untersuchungen mit dem deutschen Taxpayer-Panel“, Wirtschaft und Statistik 2/2022, S. 30-41.
- Schneider, D. (2012a): Das Steuerstatistische Gesamtsystem. Teil 1: Entstehungsgeschichte. Statistische Monatshefte Rheinland-Pfalz, 2/2012, S. 135-138. Online verfügbar unter: <https://www.statistik.rlp.de/fileadmin/dokumente/monatshefte/2012/Februar/02-2012-135.pdf> [zuletzt abgerufen am 19.11.2024]
- Schneider, D. (2012b): Das Steuerstatistische Gesamtsystem. Teil 2: Aktueller Stand der Umsetzung. Statistische Monatshefte Rheinland-Pfalz, 3/2012, S. 227-234. Online verfügbar unter: <https://www.statistik.rlp.de/fileadmin/dokumente/monatshefte/2012/Maerz/03-2012-227.pdf> [zuletzt abgerufen am 19.11.2024]
- Statistisches Bundesamt (2024): Nutzungskonzept Taxpayer-Panel 2001-2020. Online verfügbar unter: https://www.forschungsdatenzentrum.de/sites/default/files/tpp_2001-2020_on-site_nk.pdf [zuletzt abgerufen am 19.11.2024]