

Public-Use-File zur Zeitverwendungserhebung 2012/13 - Anonymisierungskonzept -

1 Einführung

Public-Use-Files sind standardisierte absolut anonymisierte Einzeldatensätze, die von den Statistischen Ämtern für jeden Interessierten bereitgestellt werden. Im Gegensatz zu Gastwissenschaftlerarbeitsplätzen oder der kontrollierten Datenfernverarbeitung werden Public-Use-Files außerhalb der geschützten Räume der amtlichen Statistik genutzt.

Absolut anonymisierte Mikrodaten fallen unter §16 Abs. 1 Pkt. 4 des BStatG und sind vom Geheimhaltungsgebot ausgenommen. Vor einer Weitergabe der Daten als Public-Use-File (PUF) in absolut anonymisierter Form muss sichergestellt werden, dass eine Zuordnung der Einzelangaben zu den Merkmalsträgern nicht mehr möglich ist.

Die Daten werden dem Nutzer nach Abschluss eines Nutzungsvertrages durch die Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder bereitgestellt.

2 Datenmaterial

Grundlage für die Erstellung des Public-Use-Files ist die Zeitverwendungserhebung 2012/2013. In den Jahren 2012/2013 hat das Statistische Bundesamt in Zusammenarbeit mit den Statistischen Ämtern der Länder zum dritten Mal nach 1991/1992 und 2001/2002 die Erhebung zur Zeitverwendung privater Haushalte durchgeführt. Die gewonnenen Daten geben Aufschluss darüber, wie Menschen in den unterschiedlichsten Lebenslagen und Haushaltskonstellationen ihre Zeit für verschiedene Lebensbereiche einteilen. Sie liefern beispielsweise Informationen zu Erwerbsarbeit und unbezahlter Arbeit, zur Arbeitsteilung innerhalb von Familien, zu Kinderbetreuung, freiwilligem Engagement sowie zur Zeitverwendung von Kindern, Jugendlichen und älteren Menschen.

In einer schriftlichen Befragung auf freiwilliger Basis gaben etwa 5000 Privathaushalte mit über 11000 Personen ab zehn Jahren zu ihrer Zeitverwendung Auskunft. Zur Ziehung der Stichprobe wurde ein Quotenverfahren genutzt. Als Quotierungsmerkmale dienten hierbei Bundesland, Haushaltstyp sowie die soziale Stellung des Haupteinkommensbeziehers im Haushalt. Um saisonale Verzerrungen zu vermeiden, fand die Befragung der gezogenen Haushalte kontinuierlich verteilt über den Zeitraum August 2012 bis Juli 2013 statt. Die gewonnenen Daten wurden anhand des Mikrozensus 2012 auf die deutsche Bevölkerung hochgerechnet.

Neben einem Haushaltsfragebogen füllte jedes Haushaltsmitglied ab zehn Jahren einen Personenfragebogen sowie ein Tagebuch aus. Der Haushaltsfragebogen enthält unter anderem Fragen zur Zusammensetzung des Haushalts, zu Wohnverhältnissen und der Inanspruchnahme von Betreuungsangeboten für Kinder unter zehn Jahren. Der Personenfragebogen deckt Angaben zu Alter und Familienstand, Erwerbstätigkeit, Einkommen, Schulbildung, Freizeitaktivitäten und freiwilligem Engagement der einzelnen Haushaltsmitglieder ab. Zudem beantworteten die Teilnehmenden Fragen zu ihrem subjektiven Zeitempfinden und zu Zeitstress in verschiedenen Lebensbereichen.

Im dritten Erhebungsteil hielten schließlich alle Personen ab zehn Jahren an drei vorgegebenen Tagen – zwei Wochentagen und einem Wochenendtag – ihren Tagesablauf fest. Dazu wurde ein Tagebuch verwendet, in dem die Tage bereits in einzelne Zeilen zu je zehn Minuten strukturiert waren. Die Befragten beschrieben hier in eigenen Worten, welcher Hauptaktivität sie jeweils nachgingen. Weiter gaben sie durch Ankreuzen vorgegebener Kategorien (Partner/Partnerin, Kinder unter zehn Jahren, andere Haushaltsmitglieder, andere bekannte Personen) an, welche Personen dabei beteiligt waren. Gegebenenfalls konnte zur gleichen Zeit eine Nebenaktivität eingetragen werden. Waren die Befragten gerade auf dem Weg von einem Ort zum anderen, war auch das genutzte Verkehrsmittel festzuhalten. Am Abschluss jedes Tages standen Fragen zur subjektiven Einschätzung des konkreten Tagesverlaufs. Die Befragten konnten hier angeben, welche Tätigkeiten ihnen die größte Freude und welche keine Freude gemacht hatten und wofür sie sich mehr Zeit gewünscht hätten.

Um eine statistische Auswertung der vielfältigen frei eingetragenen Tätigkeiten zu ermöglichen, wurden diese im Statistischen Bundesamt mit Hilfe eines einheitlichen Verzeichnisses von 165 verschiedenen Aktivitäten codiert. Dieses umfasst den persönlichen Bereich, Erwerbstätigkeit, Bildung, Haushaltsführung und Betreuung der Familie, ehrenamtliche und freiwillige Tätigkeiten, soziales Leben, Sport und Hobbys, Mediennutzung und Wegezeiten. Neben der Erfassung der Erhebungsunterlagen und Codierung der Aktivitäten wurden die Daten eingehend geprüft und plausibilisiert.

Das Aktivitätenverzeichnis wurde anhand der folgenden Aspekte erstellt. Zum einen sollten internationale Vergleiche mit Zeitverwendungserhebungen anderer europäischer Länder ermöglicht werden. Daher orientiert sich das Verzeichnis sehr an den aktuellen Empfehlungen des Statistischen Amtes der Europäischen Union, die in den „Harmonised European Time Use Surveys 2008 Guidelines“ festgehalten sind. Zum anderen sollte es so weit wie möglich kompatibel mit dem Verzeichnis der Zeitbudgeterhebung 2001/2002 sein, um Aussagen über die Veränderung der Ergebnisse über die Zeit treffen zu können. Weiter bildet die Liste Tätigkeiten, die in familien-, bildungs- und sozialpolitischer Hinsicht besonders interessant sind, möglichst detailliert ab. Schließlich enthält sie aufgrund des im Tagebuch vorgegebenen 10-Minuten-Taktes keine Aktivitäten, die grundsätzlich weniger als zehn Minuten andauern.¹

3 Anonymisierung

Die Anordnung der Datensätze des zu übermittelnden Datenbestands ist systemfrei. Direkte Identifikatoren und Hilfsmerkmale wie beispielsweise der Name und das Geburtsdatum sind im Public-Use-File nicht enthalten. Eine Übersicht über die im Public-Use-File enthaltenen Merkmale findet sich in den Datensatzbeschreibungen.

Analysen haben gezeigt, dass das Risiko der Reidentifizierung bei Haushalten mit besonders vielen Haushaltsmitgliedern erhöht ist. In Haushalten mit sieben und mehr Haushaltsmitgliedern werden deshalb ausgewählte Personen gelöscht. Bei der Auswahl der zu löschenden Personen wird darauf geachtet, dass der Haushaltstyp unverändert bleibt. Keinesfalls werden Personen gelöscht, die verheiratet sind und deren Ehepartner im Haushalt lebt, das jüngste Kind im Haushalt oder selbstständige Personen. Die Datensätze werden aus der Personen-, Zeittakt- und Summendatendatei gelöscht. Die Anzahl der Haushaltsmitglieder und die Anzahl der Kinder

¹ Vgl. Maier, Lucia, 2014: Methodik und Durchführung der Zeitverwendungserhebung 2012/2013, In: Statistisches Bundesamt (Hrsg.), Wirtschaft und Statistik, Wiesbaden, S. 672ff.

unter zehn Jahren werden entsprechend angepasst. Des Weiteren werden Haushalte mit gleichgeschlechtlichen Paaren entfernt, um die absolute Anonymität der Daten sicherzustellen. Die entsprechenden Datensätze werden ebenfalls in der Haushalts-, Personen-, Zeittakt- und Summendatendatei gelöscht.

Da der Auswahlatz der Zeitverwendungserhebung mit ca. 0,01% der Haushalte in Deutschland vergleichsweise klein ist, sind die Datenbestände schon von vornherein sehr gut vor Reidentifikationsversuchen geschützt. Der Datenangreifer weiß auch bei einer eindeutigen Zuordnung nie mit Sicherheit, ob er den richtigen Haushalt zugeordnet hat, wenn die erhobenen Daten die Grundgesamtheit nicht vollständig erfassen. Die von der Stichprobeneigenschaft ausgehende Schutzwirkung besteht jedoch dann nicht mehr, wenn dem Angreifer bekannt ist, dass der von ihm gesuchte Haushalt an der amtlichen Erhebung teilgenommen hat. Das Risiko des spezifischen Wissens ist sehr klein, aber es kann nicht ausgeschlossen werden, dass die Befragten selbst Dritten mitteilen, dass sie an der Zeitverwendungserhebung teilgenommen haben. Dieses spezifische Risiko kann gezielt verringert werden, indem nur eine nach dem Zufallsprinzip gezogene Substichprobe des ursprünglichen Datenfiles weitergegeben wird.²

Die Beschränkung der Zeitverwendungserhebungsdaten auf eine 80%-Stichprobe mit 4221 Haushalten stellt damit eine weitere Anonymisierungsmaßnahme dar. Die Unterstichprobe der Haushalte wird zufällig gezogen, nachdem die nicht berücksichtigten Haushalte entfernt worden sind. Ein Datenangreifer verliert bei einer versuchten Identifikation durch die Stichprobe die Kenntnis, ob der gesuchte Haushalt in der Stichprobe enthalten ist. Ordnet er einen Datensatz aus dem Zusatzwissen einem Haushalt zu, so trägt er das Risiko, dass diese Zuordnung nur deshalb zustande kam, weil der richtige Haushalt nicht in der Stichprobe enthalten ist. Die Zuordnung ist für ihn wertlos. Die Stichprobenziehung ist ein entscheidender Beitrag zur Erreichung der absoluten Anonymität.

Im Folgenden werden die Anonymisierungsmaßnahmen beschrieben, die auf die einzelnen Merkmale der Haushalts-, Personen-, Zeittakt- und Summendatensätze angewendet werden.

3.1 Haushaltsdatensätze

id_hhx Haushaltsnummer

Das Merkmal wird als fortlaufende Nummer der systemfreien Anordnung neu generiert.

hb9x Wohnfläche Wohnung

Die Wohnfläche der Wohnung wird bis 190 qm in 5-qm-Schritten gerundet, ab 190 qm wird in 10 qm-Schritten vergrößert. Wohnflächen unter 30 qm werden zu einer Klasse zusammengefasst. Ab 200 qm Wohnfläche wird ein Top-Coding durchgeführt.

hb10x Anzahl der Räume

In Haushalten mit acht und mehr Wohn- und Schlafräumen wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Anzahl der Räume „8 und mehr“ beträgt.

² Vgl. Müller, Walter/Blien, Uwe/Knoche, Peter/Wirth, Heike u.a., 1991: Die faktische Anonymität von Mikrodaten. In: Statistisches Bundesamt (Hrsg.), Schriftreihe Forum der Bundesstatistik, Band 19. Stuttgart: Metzler-Poeschel, S. 400ff.

hb121bx Auto Anzahl

In Haushalten mit drei und mehr Autos wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Anzahl der Autos „3 und mehr“ beträgt.

hb122bx Computer Anzahl

In Haushalten mit sechs und mehr Computern wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Anzahl der Computer „6 und mehr“ beträgt.

hc141a bis hc1411c Unterstützung des Haushalts durch andere Personen

Bei den Stundenangaben zu den einzelnen Unterstützungsleistungen wird jeweils ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Stundenzahl z.B. „60 und mehr“ beträgt. Daneben werden einige wenige Einzelfälle mit jeweils benachbarten Stundenangaben zusammengefasst.

hd15x Monatliches Haushaltsnettoeinkommen

Bei dem monatlichen Nettoeinkommen des Haushaltes werden die Einkommensklassen zu fünf Klassen zusammengefasst. In der Folge wird bei den betroffenen Fällen nur ausgewiesen, dass das monatliche Haushaltsnettoeinkommen „Unter 1100 Euro“, „1100 bis unter 1700 Euro“, „1700 bis unter 2300 Euro“, „2300 bis unter 3600 Euro“ bzw. „3600 Euro und mehr“ beträgt.

Zusätzliche Maßnahmen bei den Haushaltsdatensätzen

Als regionale Information werden die Gebietsstände West (alte Bundesländer) und Ost (neue Bundesländer einschließlich Berlin) weitergegeben. Aufgrund des geringen Stichprobenumfangs der Zeitverwendungserhebung wird das Merkmal Bundesland, Gemeindegrößenklasse und Siedlungsstruktureller Kreistyp gelöscht.

3.2 Personendatensätze***alterx Alter***

Bei Kindern unter zehn Jahren wird das Alter in Gruppen zusammengefasst (0 bis unter 3 Jahren / 3 bis unter 6 Jahren / 6 Jahre / 7 bis unter 10 Jahren). Bei Personen, die 75 Jahre und älter sind, wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Altersgruppen nur ausgewiesen, dass diese „0 bis unter 3 Jahre“, ..., bzw. „75 Jahre und älter“ sind. Eine Zufallsauswahl von 50% der Personen erhält eine veränderte Altersangabe innerhalb folgender Altersklassen: 10-13, 14-17, 18-21, 22-26, 27-34, 35-44, 45-54, 55-64, 65-77 Jahre. In besonderen Haushaltskonstellationen wurde zusätzlich das Alter einzelner Personen verändert.

pa2x Familienstand

Die Familienstände „geschieden“, „verheiratet und dauernd getrennt lebend“, „in eingetragener Lebenspartnerschaft getrennt lebend“ sowie „eingetragene Lebenspartnerschaft aufgehoben“ werden zu einer Kategorie zusammengefasst. Ebenfalls zu einer Kategorie zusammengefasst werden die beiden Familienstände „verwitwet“ und „eingetragener Lebenspartner verstorben“.

ha6x Stellung im Haushalt

Der Einzelwert „Großvater/Großmutter“ des Merkmals Stellung im Haushalt wird in „anders verwandt/verschwägert“ geändert.

ha71x Erste Staatsangehörigkeit

Die Ausprägungen „Übrige Europäische Union“ sowie „Sonstige Staatsangehörigkeit, staatenlos“ werden zusammengefasst. Das Merkmal „Zweite Staatsangehörigkeit“ wird nicht weitergegeben.

ha8x Geburtsland

Die Ausprägungen „Übrige Europäische Union“ sowie „Sonstiges Land“ werden zusammengefasst.

he18_kibetrx Kinderbetreuung (Betreute Zeit insgesamt durch Tagesmutter/Tagesvater, Krippe, Kindertagesstätte, vorschulische Einrichtungen, Hort)

Die einzelnen Betreuungsarten he182 bis he186 werden zu einem Merkmal zusammengefasst. Dabei werden die Wochenstunden der einzelnen Betreuungsarten addiert. Danach wird das Merkmal Kinderbetreuung im mittleren Wertebereich mit einer Klassenbreite von fünf Stunden und im oberen und unteren Wertebereich mit einer Klassenbreite von zehn Stunden klassifiziert. Ab 41 Stunden wird ein Top-Coding durchgeführt.

he187x Betreute Zeit durch Verwandte, Freunde, Nachbarn

Bei der betreuten Zeit durch Verwandte, Freunde, Nachbarn von 21 Stunden und mehr wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die betreute Zeit durch Verwandte, Freunde, Nachbarn „21 Stunden und mehr“ beträgt. Des Weiteren werden bei dem Merkmal Einzelwerte verändert.

he19x Betreuung Tage pro Woche

Bei der Betreuung eines Kindes unter zehn Jahren von fünf und mehr Tagen pro Woche wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Betreuung „5 Tage und mehr“ beträgt.

he202x Unterrichtsstunden pro Woche (HFB)

Das Merkmal Unterrichtsstunden pro Woche von Kindern unter zehn Jahren wird im oberen und unteren Wertebereich mit einer Klassenbreite von fünf bzw. zehn Stunden klassifiziert. Als Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Unterrichtsstunden pro Woche „15 - 20 Stunden“ bzw. „31 - 40 Stunden“ betragen.

he212x AG-Stunden pro Woche (HFB)

Bei Kindern unter zehn Jahren mit fünf und mehr AG-Stunden pro Woche wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die AG-Stunden pro Woche „5 Stunden und mehr“ betragen.

he22x Zeitstunden insgesamt in der Schule pro Woche (HFB)

Bei den wöchentlichen Zeitstunden in der Schule von Kindern unter zehn Jahren werden die oberen Werte zusammengefasst. Als Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Zeitstunden in der Schule pro Woche „41 - 50 Stunden“ betragen. Außerdem werden bei dem Merkmal Einzelwerte verändert.

he241x Stunden pro Woche: Sport und Bewegung außerhalb der Betreuungseinrichtung/Schule

Bei Kindern unter zehn Jahren, die wöchentlich sechs und mehr Stunden außerschulische Freizeitangebote im Bereich Sport und Bewegung nutzen, wird ein Top-Coding durchgeführt. In

der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass diese „6 Stunden und mehr“ pro Woche Sport- und Bewegungsangebote außerhalb der Betreuungseinrichtung/Schule nutzen.

he24_sprachfx Stunden pro Woche: Sprachförderung (Deutsch, andere Sprache) außerhalb der Betreuungseinrichtung/Schule

Aufgrund der geringen Fallzahl von Kindern unter zehn Jahren, die Sprachförderung in Deutsch, Sprachförderung in einer anderen Sprache oder Nachhilfe/Förderkurse außerhalb der Betreuungseinrichtung/Schule erhalten, und des damit verbundenen Reidentifikationsrisikos werden die Stundenangaben dieser drei Merkmale nicht einzeln weitergegeben sondern zu Variable he249x hinzugefügt.

he245x Stunden pro Woche: Singen, Musikinstrument spielen außerhalb der Betreuungseinrichtung/Schule

Bei Kindern unter zehn Jahren, die wöchentlich drei und mehr Stunden außerschulische Freizeitangebote im Bereich Singen und Musikinstrument spielen nutzen, wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass diese „3 Stunden und mehr“ pro Woche Angebote im Bereich Singen und Musikinstrument spielen außerhalb der Betreuungseinrichtung/Schule nutzen.

he246x Stunden pro Woche: Malen, Basteln außerhalb der Betreuungseinrichtung/Schule

Bei Kindern unter zehn Jahren, die wöchentlich drei und mehr Stunden außerschulische Freizeitangebote im Bereich Malen, Basteln nutzen, wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass diese „3 Stunden und mehr“ pro Woche Angebote im Bereich Malen, Basteln außerhalb der Betreuungseinrichtung/Schule nutzen.

he247x Stunden pro Woche: Tanzen, Theater spielen außerhalb der Betreuungseinrichtung/Schule

Bei Kindern unter zehn Jahren, die wöchentlich zwei und mehr Stunden außerschulische Freizeitangebote im Bereich Tanzen, Theater spielen nutzen, wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass diese „2 Stunden und mehr“ pro Woche Angebote im Bereich Tanzen, Theater spielen außerhalb der Betreuungseinrichtung/-Schule nutzen.

pc5x Berufliche Stellung in der Haupterwerbstätigkeit

Gering besetzten Merkmalsausprägungen der beruflichen Stellung werden zusammengefasst. Mithelfende Familienangehörige werden mit den Selbstständigen und Landwirten zusammengefasst, der freiwillige Wehrdienst/Bundesfreiwilligendienst und das freiwillige Jahr werden mit den Angestellten zusammengefasst.

***pc8x Wochenstunden in der Haupterwerbstätigkeit,
pc9x Wunscharbeitszeit in der Haupterwerbstätigkeit***

Angaben zwischen null und drei Stunden werden zusammengefasst. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Wochenstunden „0 bis 3 Stunden“ betragen. Bei der Wochenarbeitszeit und der Wunscharbeitszeit in der Haupterwerbstätigkeit wird bei 60 Stunden und mehr ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Wochenstunden in der Haupterwerbstätigkeit bzw. die Wunscharbeitszeit in der Haupterwerbstätigkeit „60 Stunden oder mehr“ betragen. Des Weiteren werden bei den beiden Merkmalen Einzelwerte derart umcodiert, dass sich die Relation zwischen der Wochen-

und Wunscharbeitszeit nicht ändert. Ein Beispiel: Arbeitet eine Person 47 Stunden pro Woche und gibt die Person als Wunscharbeitszeit auch 47 Stunden an und ist dieser Wert bei dem letztgenannten Merkmal ein Einzelwert, werden beide Merkmale entsprechend geändert.

pc11x Wirtschaftszweig

Der Einzelwert der Merkmalsausprägung „Konsulat, Botschaft, inter- und supranationale Organisation“ wird in einen anderen Wirtschaftszweig umcodiert. Die Wirtschaftszweige „Energieversorgung“, „Wasserversorgung, Abwasser, Abfallentsorgung“, sowie „Bergbau und Gewinnung von Erdöl, Erdgas, Steinen“ werden zusammengefasst.

pc122x Kilometer zur Arbeitsstätte in der Haupterwerbstätigkeit

Die Werte 0 km und 1 km werden zusammengefasst. Der Arbeitsweg in Kilometern wird ab einer Entfernung von mehr als 30 km in 10-km-Schritten gerundet. Ferner wird ab einer Entfernung von 90 km ein Top-Coding durchgeführt.

pc13x Arbeitsweg in Minuten in der Haupterwerbstätigkeit

Der Arbeitsweg in Minuten wird bis 80 Minuten in 5-Minuten-Schritten gerundet, bei einem Arbeitsweg bis 100 Minuten wird in 10-Minuten-Schritten vergrößert. Ab 100 Minuten wird ein Top-Coding durchgeführt.

pc16_fruehsonstx, pc16_spaetnachtx Früh-, Spät-, Nachtschicht, Sonstige Schicht

Die Nennungen „Frühschicht“ und „Sonstige Schicht“ werden zu einem Merkmal zusammengefasst. Die Nennungen „Spätschicht“ und „Nachtschicht“ werden zu einem weiteren Merkmal zusammengefasst.

pd192x Anzahl der Nebenerwerbstätigkeiten

Bei Erwerbstätigen mit zwei und mehr Nebenerwerbstätigkeiten wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Anzahl der Nebenerwerbstätigkeiten „2 oder mehr“ beträgt.

pd21x Tage pro Monat, an denen Nebenerwerbstätigkeit ausgeübt wird,

pd22x Stunden pro Tag, an denen Nebenerwerbstätigkeit ausgeübt wird

Bei diesen beiden Merkmalen werden Einzelwerte umcodiert.

pe23x Monatliches Nettoeinkommen aus Haupterwerbstätigkeit und Nebenerwerbstätigkeit(en)

Bei dem monatlichen Nettoeinkommen aus Erwerbstätigkeit werden die Einkommensklassen zu fünf Klassen zusammengefasst. In der Folge wird bei den betroffenen Fällen nur ausgewiesen, dass das monatliche Einkommen „Unter 900 Euro“, „900 bis unter 1300 Euro“, „1300 bis unter 1700 Euro“, „1700 bis unter 2300 Euro“ bzw. „2300 Euro oder mehr“ beträgt.

pf25x Art der Schule

Die Angaben zur Art der Schule werden zu zwei Kategorien (allgemeinbildende Schulen und berufliche Schulen) zusammengefasst.

pf26x Unterrichtsstunden pro Woche (PFB)

Bei Schülern, die unter acht Unterrichtsstunden pro Woche haben, wird ein Bottom-Coding durchgeführt, der obere Wertebereich von 41 bis 45 Unterrichtsstunden wird zu einer Klasse zusammengefasst. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die

Unterrichtsstunden pro Woche „unter 8 Stunden“ bzw. „41 bis 45 Stunden“ betragen. Ferner werden bei dem Merkmal Einzelwerte umcodiert.

pf272x AG-Stunden pro Woche (PFB)

Bei Schülern mit fünf und mehr AG-Stunden pro Woche wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die AG-Stunden pro Woche „5 Stunden und mehr“ betragen.

pf28x Zeitstunden insgesamt in der Schule pro Woche (PFB)

Bei Schülern, die wöchentlich weniger als acht Zeitstunden in der Schule verbringen, wird ein Bottom-Coding, vorgenommen. Bei Schülern, die 45 Zeitstunden oder mehr in der Schule pro Woche verbringen, wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Personen nur ausgewiesen, dass diese wöchentlich „unter 8 Stunden“ oder „45 Stunden und mehr“ in der Schule verbringen. Des Weiteren werden bei dem Merkmal Einzelwerte umcodiert.

pf291x Sport außerhalb der Schule

Bei Personen mit elf und mehr Stunden Sport wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass der Zeitaufwand für Sport außerhalb der Schule „11 Stunden und mehr“ beträgt.

***pf292x Sprachförderung in Deutsch außerhalb der Schule,
pf293x Sprachförderung andere Sprache außerhalb der Schule***

Bei Personen mit zwei und mehr Stunden Sprachförderung in Deutsch bzw. in einer anderen Sprache wird ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Sprachförderung in Deutsch bzw. in einer anderen Sprache außerhalb der Schule „2 Stunden und mehr“ beträgt.

pf294x Nachhilfe, Förderkurse außerhalb der Schule

Bei Schülern mit vier und mehr Stunden Nachhilfe, Förderkurse wird ebenfalls ein Top-Coding durchgeführt. In der Folge wird bei den betreffenden Fällen nur ausgewiesen, dass die Nachhilfe, Förderkurse außerhalb der Schule „4 Stunden und mehr“ beträgt.

pf31_typ Höchster beruflicher Ausbildungs- oder Hochschulabschluss

Die Angaben zum höchsten beruflichen Ausbildungs- oder Hochschulabschluss werden wie folgt zusammengefasst:

- Berufsvorbereitungsjahr, Anlernausbildung und berufliches Praktikum
- Lehre, Berufsausbildung, Vorbereitungsdienst für den mittleren Dienst in der öffentlichen Verwaltung, Berufsfachschule/Kollegscheule und 1-jährige Schule des Gesundheitswesens
- 2- oder 3-jährigen Schule des Gesundheitswesens, Meister, Techniker oder gleichwertiger Fachschulabschluss und Fachakademie
- Berufsakademie, Verwaltungsfachhochschule, Fachhochschule (auch Ingenieurschule), Universität, wissenschaftliche Hochschule, Kunsthochschule und Promotion

pj37 Unterstützung für Haushaltsmitglieder

Die Angaben zur Unterstützung von Haushaltsmitgliedern, die Leistungen aus der Pflegeversicherung erhalten, werden nicht weitergegeben.

pk402ax bis pk402ex Aktivitäten, für die Befragte sich mehr Zeit wünschen

Einige Aktivitätscodes werden umcodiert bzw. zusammengefasst. Vergleiche dazu die Ausführungen zu Zeittakt- und Summendatensätzen unter 3.3.

Zusätzliche Maßnahmen bei den Personendatensätzen

Bei einigen Datensätzen müssen zusätzliche Maßnahmen durchgeführt werden. Durch die Kombination ausgewählter Merkmale kann ein erhöhtes Risiko für die Identifikation eines Haushaltes ausgehen. Beispiele hierfür sind z.B. eine große Anzahl an Haushaltsmitgliedern und das Alter der Haushaltsmitglieder, Haushalte mit auffällig älteren Kindern, Haushalte mit besonders vielen Kindern unter zehn Jahren oder Haushalte, in denen nur Großeltern mit Enkeln leben. Betrifft die Merkmalskonstellationen nur wenige Haushalte, werden die Datensätze geringfügig verändert.

Bei besonders alten Menschen, die in großen Haushalten leben, besteht beispielsweise ein erhöhtes Reidentifikationsrisiko. Um die absolute Anonymität der Daten zu erreichen, wird in diesen Haushalten das Alter dieser Personen durch ein zufällig gewähltes Alter innerhalb einer Altersklasse ersetzt. Auch wird das Alter bei Familien mit auffällig älteren Kindern modifiziert. Leben viele Kinder unter zehn Jahren im Haushalt, wird das Alter mehrerer Personen im Haushalt geändert. Außerdem wird das Alter von Personen in besonders großen Haushalten, in denen ein Haushaltsmitglied Leistungen aus der Pflegeversicherung erhält, verändert. In Haushalten mit Enkeln wird das Alter des ältesten Enkels ebenfalls geringfügig verändert.

3.3 Zeittakt- und Summendatensätze**Zusätzliche Maßnahmen bei den Zeittakt- und Summendatensätzen**

Das Merkmal „Ausfülldatum“ wird gelöscht. Stattdessen wird das Kalenderquartal und das Jahr des Ausfülldatums im Public-Use-File weitergegeben.

Einige der Aktivitätscodes wurden zusammengefasst:

Zeitangaben zu Aktivität 316 (Haushalts- und Ernährungswissenschaften) wurden als 319 (Sonstige Unterrichtsfächer) codiert.

Angaben zu den Aktivitäten 331 bis 339 (verschiedene Arbeitsgemeinschaften in der Schule) werden zusammengefasst im Dreisteller 330 ausgewiesen.

Angaben zu den Aktivitäten 352 (Private Hausaufgabenbetreuung/Nachhilfe für die Schule) und 351 (Hausaufgabe/Selbstlernen für die Schule) werden zusammengefasst im Dreisteller 354 ausgewiesen.

Angaben zu den Aktivitäten 714 (Wintersport) und 718 (Kampf- und Kraftsport) wurden als 719 (Sonstige körperliche Bewegung im Bereich Sport- und Outdoor-Aktivitäten) codiert.

Angaben zur Aktivität 720 (Jagen/Fischen und Beeren/Pilze/Kräuter sammeln) wurden als 790 (Sport ohne nähere Bezeichnung) codiert.

Angaben zur Aktivität 751 (Briefmarken, Münzen, etc. sammeln) wurden als 759 (Sonstige technische und andere Hobbys) codiert.

4 Hinweise zum Vergleich mit früheren Jahren

In diesem Abschnitt werden wichtige Unterschiede der Public-Use-Files von 2012/2013 und 2001/2002 beschrieben. Die Vergleichbarkeit der Public-Use-Files von 2001/02 und 1991/1992 wurde bereits andernorts dokumentiert.³ Die Erhebungen 2012/2013 und 2001/2002 unterscheiden sich, beispielsweise in den erhobenen Merkmalen, den Merkmalsausprägungen und den Aktivitätscodes. Die Unterschiede beruhen auf geänderten Empfehlungen zur Harmonisierung der europäischen Zeitverwendungserhebungen seitens EUROSTAT, vor allem aber auf anderen Schwerpunkten in familien-, bildungs- und sozialpolitischen Fragestellungen. Zur Sicherstellung einer hohen Datenqualität flossen auch die Erfahrungen aus anderen Haushaltserhebungen und Pretest-Ergebnisse in die Konzeption der Zeitverwendungserhebung 2012/2013 ein.

Die Erhebungskonzepte der Zeitverwendungserhebungen unterscheiden sich daher zum Teil deutlich. Während 2012/2013 den Nichterwerbstätigen über die Selbsteinschätzung die soziale Stellung in der gewünschten Untergliederung eindeutig zugeordnet werden kann, ist das in der Erhebung 2001/2002 nicht möglich, da personenbezogene Angaben zur Selbsteinschätzung der sozialen Stellung nicht erfasst wurden. Entsprechende Abgrenzungen lassen sich zwar aus den Erhebungsmerkmalen der Zeitverwendungserhebung 2001/2002 konstituieren, sind aber mit 2012/2013 nicht vergleichbar.

Die Erwerbsbeteiligung ist in beiden Erhebungen an das Erwerbskonzept des Mikrozensus angelehnt. Erwerbstätig sind Personen ab 15 Jahren, deren Erwerbstätigkeit normalerweise mindestens eine Stunde pro Woche beträgt. Methodische Unterschiede gibt es aber bei der Abfrage zur Erwerbstätigkeit. In der Erhebung 2001/2002 wurden beispielweise für geringfügig Erwerbstätige eine Vielzahl an Merkmalen zur Erwerbstätigkeit nicht erhoben, die in der Zeitverwendungserhebung 2012/2013 infolge einer anderen Abgrenzung von Haupt- und Nebentätigkeit vorliegen.

Der Haushaltsfragebogen 2001/2002 enthält eine Beziehungsmatrix, mit deren Hilfe sich alle Beziehungen der Haushaltsmitglieder untereinander bestimmen lassen. In schriftlichen Befragungen bereitete den Auskunftsgibenden das Ausfüllen dieser komplexen Beziehungsmatrix große Schwierigkeiten. In der Erhebung 2012/2013 wurde deshalb nur die Beziehung zum Haupteinkommensbezieher erfragt. Diese Variablen sind jeweils in der Personendatei gespeichert. In der Personendatei sind auch die Datensätze zu den Betreuungs- und Bildungsangeboten, die von Kindern unter zehn Jahren in Anspruch genommen werden, angelegt. Außerdem sind dort die soziodemographischen Merkmale der Personen, die nicht geantwortet haben, gespeichert. Die Variablen für Kinder unter zehn Jahren sind in der Zeitverwendungserhebung 2012/2013 mit „trifft nicht zu“ und die Variablen der Personen ohne ausgefüllten Fragebogen mit „keine Angabe“ codiert. Im Public-Use-File 2001/2002 sind die Variablen für Kinder unter zehn Jahren mit „Kind unter 10 Jahren“ codiert.

Die Datensätze der Haushalts- und Personendatei enthalten zusätzliche Standardtypisierungen, die für einheitliche Definitionen für Haushalts- und Personentypen verwendet werden können. Aufgrund der verschiedenen Erhebungskonzepte sind die Typisierungen 2001/2002 und 2012/2013 nur eingeschränkt miteinander vergleichbar. Neben den bisher erörterten Unterschieden bei der Erwerbstätigkeit und dem sozialen Status wurde die Haushaltstypisierung

³ Vgl. Holz, Erlend, 2007: Vademekum Public Use Files - Zeitbudgeterhebungen (ZBE) 2001/02 und 1991/92, Statistisches Bundesamt, Bonn, S. 5ff.

anders abgegrenzt. In 2001/2002 werden als Kinder ledige Personen ohne Altersbegrenzung betrachtet, die Kind der Bezugsperson des Haushaltes und/oder des Partners sind. Bei den Standardtypisierungen zu den Haushaltstypen 2012/2013 gibt es hingegen eine Altersbegrenzung. In einem Alleinerziehenden-Haushalt oder Paar-Haushalt mit Kind gibt es mindestens ein Kind/Enkelkind unter 18 Jahren, weitere Kinder sind ledig und unter 27 Jahren. Einbezogen sind dabei grundsätzlich alle Paare, unabhängig davon, ob diese verheiratet sind oder nicht.

Wie bereits erwähnt, unterscheiden sich auch die Aktivitätscodes in beiden Erhebungen. Es wurde eine Datei gefertigt, die einen Umsteigeschlüssel für die Aktivitäten der Zeitverwendungserhebung 2012/2013 und 2001/2002 enthält und Zeitvergleiche ermöglicht. Lebenslanges Lernen stand 2001/2002 im Focus bildungspolitischer Fragestellungen. Deshalb wurden die Aktivitäten Qualifizierung/Weiterbildung für den Beruf während der Arbeitszeit (Code 241), Vor- und Nachbereitungen von Lehrveranstaltungen (Code 35) und die Qualifikation/Fortbildung außerhalb der Arbeitszeit (Code 361) in der früheren Erhebung sehr differenziert nach Methoden und Medien erfasst. Einige Aktivitäten der Zeitverwendungserhebung 2001/2002 können dabei nur einem 2-Steller der Aktivitätenliste 2012/2013 zugeordnet werden. In der aktuellen Erhebung lag der Schwerpunkt hingegen auf der detaillierten Erfassung schulischer Bildungsangebote, wie den einzelnen Unterrichtsfächern, Arbeitsgemeinschaften und Betreuungsangeboten in der Schule.

Im Bereich der Unterstützungsleistungen ist die Zeitverwendungserhebung 2012/2013 weniger detailliert. In der Erhebung 2001/2002 wurde die Pflege und Betreuung von kranken und älteren erwachsenen Familienangehörigen gesondert innerhalb der Unterstützung von erwachsenen Haushaltsmitgliedern ausgewiesen. Die informelle Hilfe für andere Haushalte wies 16 Unterkategorien auf. Aus methodischen Gründen werden die Aktivitäten Unterstützung, Pflege, Betreuung von erwachsenen Haushaltsmitgliedern (Code 48) und Unterstützung für andere Haushalte (Code 52) in der Erhebung 2012/2013 nicht tiefer untergliedert. Die Differenzierung der verschiedenen Unterstützungsleistungen wurde in einem Pretest getestet. Dabei zeigte sich, dass sich Aktivitäten wie die Alten- und Krankenpflege nicht von anderen im Haushalt anfallenden Tätigkeiten trennen lässt, wie z.B. das Zubereiten von speziellen Mahlzeiten. Außerdem sind die Aktivitäten der körperlichen Bewegung (Code 71) 2012/2013 weniger stark untergliedert als in der Zeitverwendungserhebung 2001/2002.

Ein weiterer konzeptioneller Unterschied betrifft die Angabe, mit wem die Zeit verbracht wurde. Im Gegensatz zu 2001/2002 beschränkt sich in der aktuellen Erhebung die Information, dass ein Kind unter zehn Jahren mit dabei war, auf Kinder, die im selben Haushalt wohnen. 2001/2002 war es gleichgültig, ob die Kinder unter zehn Jahren in dem Haushalt der Person, die das Tagebuch ausfüllt, leben oder nicht. Entsprechendes gilt für den Ehe-/Lebenspartner. In der Zeitverwendungserhebung 2012/2013 wurden Partner und Kinder, mit denen die Person nicht zusammen wohnt, als andere bekannte Personen eingetragen. Die Variable zum Ort der Aktivität, die 2001/2002 aus den Angaben zur Aktivität in Verbindung mit den Wegezeiten angelegt wurde, steht in der Erhebung 2012/2013 nicht zur Verfügung.

5 Fazit

Die unter Abschnitt 3 beschriebenen Anonymisierungsmaßnahmen führen zu Mikrodatenfiles, bei denen eine De-Anonymisierung einzelner Merkmalsträger ausgeschlossen ist. Die Datensätze

der Zeitverwendungserhebung 2012/2013 sind damit absolut anonym und können in dieser Form als Public-Use-Files veröffentlicht werden. Damit ist es gelungen, nach den Zeitbudgeterhebungen 1991/1992 und 2001/2002 wieder einen aktuellen absolut anonymisierten Mikrodatsatz zur Zeitverwendung in Deutschland bereitzustellen.

Literatur

Holz, Erhard, 2007: Vademekum Public Use Files - Zeitbudgeterhebungen (ZBE) 2001/02 und 1991/92, Statistisches Bundesamt, Bonn

Maier, Lucia, 2014: Methodik und Durchführung der Zeitverwendungserhebung 2012/2013, In: Statistisches Bundesamt (Hrsg.), Wirtschaft und Statistik, Wiesbaden

Müller, Walter/Blien, Uwe/Knoche, Peter/Wirth, Heike u.a., 1991: Die faktische Anonymität von Mikrodaten. In: Statistisches Bundesamt (Hrsg.), Schriftreihe Forum der Bundesstatistik, Band 19. Stuttgart: Metzler-Poeschel