

Thomas Körner, M. A., Dr. Anja Nimmergut, Jens Nökel (Statistisches Landesamt Baden-Württemberg),
Dipl.-Mathematikerin Sandra Rohloff

Die Dauerstichprobe befragungsbereiter Haushalte

Die neue Auswahlgrundlage für freiwillige Haushaltsbefragungen

Zur Durchführung freiwilliger Haushaltsbefragungen bauen die Statistischen Ämter des Bundes und der Länder seit dem Jahr 2004 eine Dauerstichprobe befragungsbereiter Haushalte auf. Die Dauerstichprobe setzt sich zusammen aus Haushalten, die sich im Anschluss an ihre Teilnahme beim Mikrozensus entschieden haben, von Zeit zu Zeit auf freiwilliger Basis an Erhebungen der amtlichen Statistik teilzunehmen. Durch den Einsatz der Dauerstichprobe können Ad-hoc-Erhebungen – auch von Teilpopulationen – sehr flexibel als Zufallsstichprobe durchgeführt werden. Zugleich werden wesentlich höhere Ausschöpfungsquoten aus der Dauerstichprobe erzielt als bei vergleichbaren Erhebungen mit herkömmlichen Verfahren der Stichprobenziehung. Dieser Beitrag stellt die Rolle der Dauerstichprobe im System der Haushaltsstatistiken dar und beschreibt die methodische Konzeption sowie künftige Einsatzmöglichkeiten.

1 Die Dauerstichprobe im System der Haushaltsstatistiken

Vor dem Hintergrund neuer sozial- und wirtschaftspolitischer Herausforderungen lässt sich seit einiger Zeit ein erhöhter Bedarf an aktuellen und verlässlichen statistischen Informationen feststellen. Zu zahlreichen Themen können Haushalts- und Personenbefragungen der amtlichen Statistik wichtige Beiträge liefern. Hierzu zählen zum Beispiel Fragen, wie wirksame Reaktionen auf den demografischen Wandel gestaltet werden müssen, wie Armut und soziale Ausgrenzung bekämpft werden können oder wie die

Chancengleichheit bei der Bildung gefördert werden kann. Mit der Dauerstichprobe befragungsbereiter Haushalte steht der amtlichen Statistik heute ein Instrument zur Verfügung, welches es ermöglicht, Haushalts- und Personenerhebungen flexibel, aktuell und mit hoher Genauigkeit durchzuführen.

Der Aufbau der Dauerstichprobe geht zurück auf die Anforderung, die neue Europäische Statistik über Einkommen und Lebensbedingungen (EU-SILC) auf Basis einer Zufallsstichprobe durchzuführen. Basierend auf den positiven Ergebnissen einer umfangreichen Pilotstudie, die von 2000 bis 2003 durchgeführt wurde¹⁾, haben die Leiterinnen und Leiter der Statistischen Ämter des Bundes und der Länder im November 2003 beschlossen, für die Durchführung von EU-SILC die Dauerstichprobe befragungsbereiter Haushalte bundesweit aufzubauen und hierzu Haushalte im Anschluss an das letzte Mikrozensusinterview um ihre Teilnahme zu bitten. Obwohl der Ausgangspunkt zur Schaffung der Dauerstichprobe die Durchführung von EU-SILC war, sieht bereits der Beschluss der Amtsleiterinnen und Amtsleiter vor, dass „die Dauerstichprobe zusätzlich für weitere Erhebungen genutzt werden soll, soweit die Erhebungsziele von EU-SILC nicht gefährdet werden.“²⁾ Hiermit wird nicht zuletzt dem Umstand Rechnung getragen, dass das Potenzial der Dauerstichprobe nur dann voll ausgeschöpft werden kann, wenn sie als Basis für mehrere Erhebungen Anwendung findet. Zugleich steigt die Wirtschaftlichkeit der Dauerstichprobe, je mehr Erhebungen auf ihrer Grundlage durchgeführt werden. Schon während der Durchführung der Pilotstudie

1) Siehe Körner, T./Nimmergut, A.: „Zu den Möglichkeiten der Nutzung einer Dauerstichprobe befragungsbereiter Haushalte in der amtlichen Statistik“ in WiSta 5/2003, S. 391 ff.

2) Bericht über die Amtsleiterkonferenz am 11. und 12. November 2003, Statistisches Bundesamt, Wiesbaden, S. 14.

stießen die Möglichkeiten der Dauerstichprobe auf reges Interesse, was eine ganze Reihe Anfragen seitens der Fachabteilungen der statistischen Ämter wie oberster Bundesbehörden belegt.

Dieser Beitrag stellt nach einem kurzen Überblick über die Grundidee der Dauerstichprobe die verschiedenen Bausteine des Gesamtkonzeptes der Dauerstichprobe dar (Kapitel 2). Hierbei wird zunächst auf die Methodik bei der Anwerbung von Haushalten eingegangen. Im Anschluss daran werden die Rolle der Profilm Merkmale sowie die Vorgehensweise bei der Pflege der Haushalte, die sich für eine Teilnahme bei der Dauerstichprobe entschieden haben, beschrieben. Schließlich werden ein Überblick über Aufbau und Funktion der für die Dauerstichprobe entwickelten Datenbank gegeben und die für Hochrechnung und Fehlerrechnung eingesetzten Methoden und Verfahren skizziert. Kapitel 3 gibt einen Ausblick auf die Nutzungsmöglichkeiten der Dauerstichprobe für aktuelle und künftige Personen- und Haushaltsstichproben. Hierzu werden die bisherigen Anwerbungsergebnisse und die Erhebungen, die die Dauerstichprobe derzeit nutzen, dargestellt. Schließlich werden Kriterien für die Möglichkeit einer Nutzung der Dauerstichprobe durch weitere Erhebungen benannt.

1.1 Die Grundidee der Dauerstichprobe befragungsbereiter Haushalte

Die Dauerstichprobe befragungsbereiter Haushalte ist eine Auswahlgrundlage für freiwillige Befragungen der amtlichen Statistik bei Haushalten und Personen. Sie enthält Haushalte, die sich zur regelmäßigen Teilnahme an freiwilligen Erhebungen der amtlichen Statistik bereit erklärt haben. Fundament der Dauerstichprobe ist der Mikrozensus. Da die Haushalte für die Dauerstichprobe ausschließlich aus dem Mikrozensus angeworben werden, kann die Dauerstichprobe für die Ziehung von mehrstufigen Zufallsstichproben verwendet werden. Die Nutzung des Mikrozensus ist der entscheidende Vorteil für die Gewinnung hochwertiger Stichproben und zeichnet die Dauerstichprobe gegenüber vielen in der kommerziellen Markt- und Meinungsforschung genutzten „Access Panels“ aus. Die Nutzung des Mikrozensus gewährleistet, dass nur Haushalte in die Dauerstichprobe gelangen können, die nach dem Zufallsverfahren aus-

gewählt worden sind. Darüber hinaus werden detaillierte Analysen des auftretenden Nonresponse-Bias (Verzerrung durch Antwortausfälle) ermöglicht.

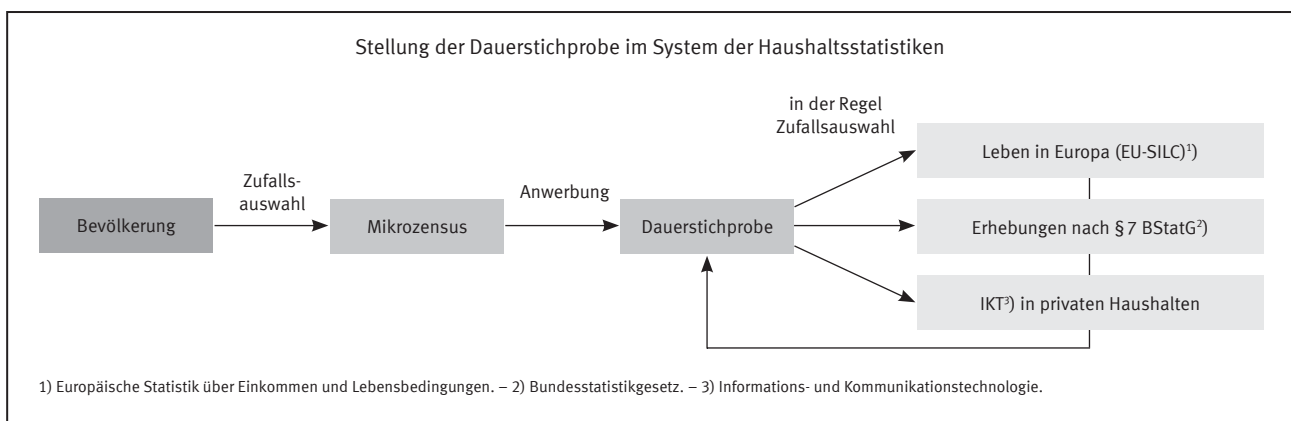
Der Mikrozensus ist die größte Haushaltsstichprobe in Deutschland und zugleich die einzige mit Auskunftspflicht. Die Stichprobe des Mikrozensus umfasst derzeit etwa 380 000 Haushalte, die nach dem Zufallsprinzip ausgewählt werden. Sie ist in Rotationsviertel mit jeweils etwa 95 000 Haushalten aufgeteilt. Jedes Rotationsviertel stellt für sich genommen eine repräsentative Stichprobe der in Deutschland lebenden wohnberechtigten Bevölkerung dar. Da durch die Auskunftspflicht nur in sehr geringem Maße Antwortausfälle zu verzeichnen sind, eignet sich der Mikrozensus in besonderer Weise als Grundlage für die Anwerbung zur Dauerstichprobe.

Auf eine mögliche Teilnahme an der Dauerstichprobe werden alle Haushalte angesprochen, die die Mikrozensus-Befragung beendet haben (siehe Abschnitt 2.1). Durch die Freiwilligkeit der Teilnahme an der Dauerstichprobe ist eine gewisse Verzerrung (Bias) unvermeidlich. Die Nutzung der Dauerstichprobe setzt daher den Einsatz wirksamer Gewichtungsverfahren voraus (siehe Abschnitt 2.5).

Mit dem Einverständnis der Haushalte werden ausgewählte sozio-demografische Variablen aus dem letzten Mikrozensusinterview in die Datenbank der Dauerstichprobe übernommen (siehe Abschnitt 2.2). Diese so genannten Profilm Merkmale werden auf einem aktuellen Stand gehalten und stehen zum Beispiel als Schichtungsvariablen zur Verfügung.

Der Kontakt zu den Haushalten wird über Erhebungen, Aktualisierungen sowie einen Newsletter mindestens einmal je Halbjahr gewährleistet, um die Haushalte laufend über die Erhebungen aus der Dauerstichprobe zu informieren und zugleich den Adressbestand möglichst aktuell zu halten. Profilm Merkmale und Adressen werden in einer komplexen Datenbankanwendung gepflegt, die das Statistische Landesamt Baden-Württemberg für den Einsatz in allen Statistischen Landesämtern entwickelt hat (siehe Abschnitt 2.4). Die Datenbank enthält zudem Informationen über die Teilnahme an Erhebungen und spielt eine wichtige Rolle bei der

Schaubild 1



Begrenzung der Belastung der Befragten sowie bei der Koordinierung der Teilnahme an Erhebungen.

1.2 Einsatzmöglichkeiten im System der Haushaltsstichproben

Die Dauerstichprobe wird im System der Haushaltsstichproben auf vielfältige Weise genutzt. Obwohl das oberste Ziel die Bereitstellung einer Zufallsstichprobe für die Statistik über Einkommen und Lebensbedingungen (EU-SILC) ist, bietet sie darüber hinaus erhebliches Potenzial für die Verbesserung des Informationsangebots über private Haushalte in Deutschland sowie für die Weiterentwicklung der Haushaltsstichproben in der amtlichen Statistik.

Die Dauerstichprobe kann für freiwillige Haushaltserhebungen genutzt werden, sofern ein Belastungsniveau der Befragten nicht überschritten wird, das die Teilnahmebereitschaft der Haushalte gefährden würde. Umgekehrt sollte die Teilnahmebereitschaft der Haushalte auch nicht durch eine „Unterforderung“ gefährdet werden. Die Entscheidung, sich an der Dauerstichprobe zu beteiligen, wird bewusst getroffen. Wenn die Haushalte nicht von Zeit zu Zeit kontaktiert werden, wird ihre Erwartung enttäuscht, an der Dauerstichprobe mitwirken zu können, oder die Entscheidung zur Teilnahme gerät schlicht in Vergessenheit.³⁾ Expertenschätzungen gehen davon aus, dass die Zahl der Ausfälle minimiert werden kann, wenn jeder Haushalt pro Jahr an drei bis vier Erhebungen teilnimmt. Auch wenn die Ausfälle der Dauerstichprobe auf einen Zusammenhang mit der Belastungsintensität weiterhin sorgfältig untersucht werden müssen, deuten die bisherigen Erfahrungen darauf hin, dass die Dauerstichprobe über EU-SILC hinaus – in Abhängigkeit von der Belastungsintensität der einzelnen Erhebungen – zumindest für zwei bis drei weitere Erhebungen je Jahr genutzt werden kann.

Erst die Nutzung der Dauerstichprobe als Standardinstrument für verschiedene Erhebungen ermöglicht es, ihr Potenzial in vollem Umfang auszuschöpfen. Die Vorteile für das System der Haushaltsstatistiken betreffen (1) die Möglichkeiten der Standardisierung der unterschiedlichen freiwilligen Haushaltserhebungen, (2) die weitere konzeptionelle Harmonisierung der Haushaltserhebungen sowie (3) die Flexibilisierung des Systems im Hinblick etwa auf kurzfristig auftretenden Datenbedarf von Bundesministerien.

(1) Zunächst ermöglicht die Dauerstichprobe eine weitgehende Standardisierung der eingesetzten Methoden und Verfahren, was zu kürzeren Vorbereitungszeiten für Erhebungen führt und den Aufwand reduziert. So können mit Hilfe der Dauerstichprobe zahlreiche Erhebungen als Zufallsstichprobe durchgeführt werden, für die dies bisher nicht möglich war. Der mit der Gewinnung von Zufallsstichproben verbundene Aufwand hat bislang dazu geführt, dass einzelne Erhebungen der amtlichen Statistik mit freiwilliger Auskunftserteilung als Quotenauswahl realisiert werden. Ähnliches gilt für die Verfahren zur Hoch- und Fehlerrechnung, die nicht

mehr für jede Statistik neu entwickelt werden müssen, sondern für alle auf Basis der Dauerstichprobe durchgeführten Erhebungen einheitlich angewendet werden können. Hierdurch können sie sehr sorgfältig und unter Abwägung zahlreicher Alternativen optimiert werden. Aktuelles Beispiel ist die Fehlerrechnung, die im Rahmen eines Kooperationsprojektes mit den Universitäten Tübingen und Trier unter der Leitung von Prof. Ralf Münnich entwickelt wird. Im Rahmen einer Simulationsstudie konnte eine Vielzahl unterschiedlicher Verfahren evaluiert und auf den Einsatz in der amtlichen Statistik zugeschnitten werden. Schließlich verspricht auch der Einsatz der Datenbank der Dauerstichprobe erhebliche Synergieeffekte. Die Verwaltung von Adressen und Profilmernken erfolgt in einem integrierten System; zugleich bietet die Datenbank Funktionalitäten für das Erhebungsmanagement sowie für die Feldsteuerung bei kleineren Erhebungen. Beispielsweise kann für die Erhebung demografischer Grunddaten zur Geburtenentwicklung, die im September 2006 durchgeführt wird, die Feldarbeit vollständig mit der Dauerstichproben-Datenbank gesteuert werden, ohne dass zusätzlicher Programmieraufwand entsteht.

(2) Neben der Standardisierung trägt die Dauerstichprobe auch zu einer fachlichen Harmonisierung der Haushaltserhebungen bei. Sie schafft eine Verbindung zwischen der einzigen Haushaltsstichprobe mit Auskunftspflicht, dem Mikrozensus, und den wesentlich kleineren freiwilligen Haushaltserhebungen. Durch die Dauerstichprobe werden erstmals die Möglichkeiten des Mikrozensus auch für freiwillige Haushaltserhebungen ausgeschöpft. Dies eröffnet mittelfristig die Möglichkeit, den Mikrozensus von Merkmalen zu entlasten, wenn zum Beispiel Merkmale, die nicht zwingend in tiefer fachlicher und regionaler Gliederung benötigt werden, verstärkt durch freiwillige Erhebungen abgedeckt werden können. In diesem Sinne könnte die Dauerstichprobe zukünftig als „Haushaltsregister“ für freiwillige Befragungen genutzt werden.

Die Dauerstichprobe leistet zugleich Vorarbeiten zur Harmonisierung der sozio-demografischen Variablen in den Haushaltsstichproben der amtlichen Statistik. Die für die Dauerstichprobe wichtigen Profilmernken werden in der Regel im Rahmen von Erhebungen aktualisiert, die auf Basis der Dauerstichprobe durchgeführt werden. Daher sollten die betreffenden Merkmale möglichst in vollem Umfang harmonisiert erhoben werden. Für die Kohärenz des Systems der Haushaltsstatistiken wäre dies ein entscheidender Vorteil.

(3) Schließlich ermöglicht die Dauerstichprobe, flexibler auf kurzfristig auftretenden Informationsbedarf zu reagieren. Die weitgehend standardisierten Verfahren ermöglichen die flexible Durchführung von Ad-hoc-Erhebungen, etwa für aktuelle politische Fragestellungen. Die stets aktuell in der Datenbank vorliegenden Profilmernken ermöglichen überdies, Befragungen ohne vor-

³⁾ Siehe Nimmergut, A./Meyer, I./Körner, T.: „Haushalte Heute. Pilotstudie zur Umsetzbarkeit einer Dauerstichprobe befragungsbereiter Haushalte in der amtlichen Statistik“, Wiesbaden 2004, S. 56 ff.

herige Abfrage von Schichtungsmerkmalen durchzuführen. In besonderem Maße ist die Dauerstichprobe für die Befragung von Teilpopulationen geeignet. So können auf Basis der Profilmkmale bestimmte Personenkreise, wie etwa Erwerbstätige, junge Erwachsene, allein Erziehende oder Senioren, direkt angesprochen werden, ohne dass zuvor ein aufwändiges Screening erforderlich ist. Dies reduziert zudem den Aufwand der Feldarbeit und trägt auch zu einer zeitnahen Bereitstellung der Ergebnisse bei.

Schließlich kann der Aufwand der Feldarbeit zusätzlich durch die gegenüber herkömmlichen Zufallsstichproben mit freiwilliger Auskunftserteilung erheblich höheren Ausschöpfungsquoten reduziert werden. Um einen vorgegebenen Nettostichprobenumfang zu erreichen, genügen wesentlich kleinere Bruttostichprobenumfänge aus der Dauerstichprobe, was nicht nur den Erhebungsaufwand reduziert, sondern auch die Feldphase verkürzt und somit auch die Aktualität der Ergebnisse verbessern kann. Darüber hinaus wird unnötiger Ärger mit Personen vermieden, die nicht zur Auskunftserteilung bereit sind.

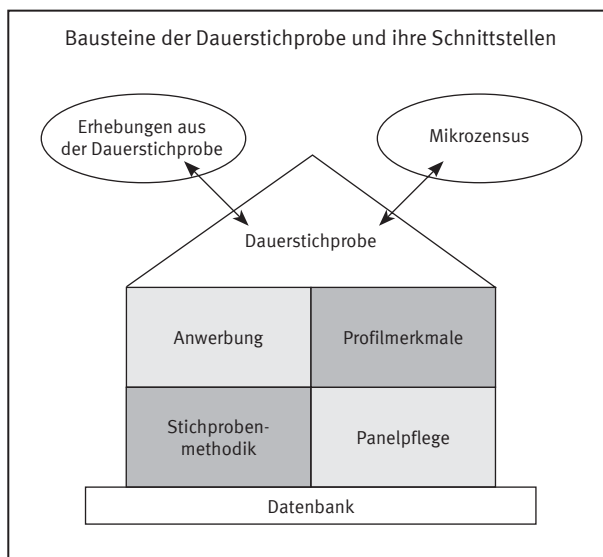
2 Gesamtkonzept der Dauerstichprobe

Die Implementierung der Dauerstichprobe hat zahlreiche fachliche, methodische, organisatorische und nicht zuletzt auch technische Aspekte und muss gleichzeitig vielfältigen Anforderungen Rechnung tragen. Sie wird aufgebaut, um eine flexibel nutzbare Auswahlgrundlage für die Durchführung freiwilliger Haushaltserhebungen zu unterschiedlichen Themenbereichen vorhalten zu können. Die Teilnahmequoten an Erhebungen aus dieser Auswahlgrundlage sollen gegenüber herkömmlichen Stichprobenverfahren deutlich höher sein. Die Dauerstichprobe muss dem Anspruch gerecht werden, möglichst effizient aufgebaut und gepflegt werden zu können und zugleich die Gegebenheiten im föderalen System bei jedem Arbeitsschritt zu berücksichtigen. Um diesen vielfältigen Anforderungen angemessen Rechnung zu tragen, wurden Konzepte entwickelt, die alle notwendigen Arbeitsschritte der Dauerstichprobe beschreiben und die verschiedenen Schnittstellen, die bei der Durchführung der einzelnen Arbeitsschritte berücksichtigt werden müssen, definieren.

Die für Aufbau, Pflege und Nutzung der Dauerstichprobe erforderlichen Arbeitsschritte können in fünf Bausteine untergliedert werden, die in Schaubild 2 dargestellt sind. Der Erfolg der Dauerstichprobe und die Güte der aus der Dauerstichprobe gezogenen Stichproben hängen von jedem der fünf Bausteine und ihrem funktionierenden Zusammenspiel ab.

An den Stellen, an denen die Bausteine in Verbindung mit anderen Statistiken stehen, ergeben sich so genannte Schnittstellen. Nur ein möglichst reibungsloser Datenfluss an diesen Schnittstellen zwischen den beteiligten Statistiken und der Dauerstichprobe ermöglicht ein effektives Funktionieren des jeweiligen Bausteins. Die Anwerbung der Dauerstichprobenteilnehmer aus dem Mikrozensus, die

Schaubild 2



Nutzung der Dauerstichprobe für die Stichprobenziehung unterschiedlicher Statistiken sowie die Aktualisierung der Profilmkmale führen zu mannigfachen Datenflüssen zwischen der Dauerstichprobe und anderen Statistiken.

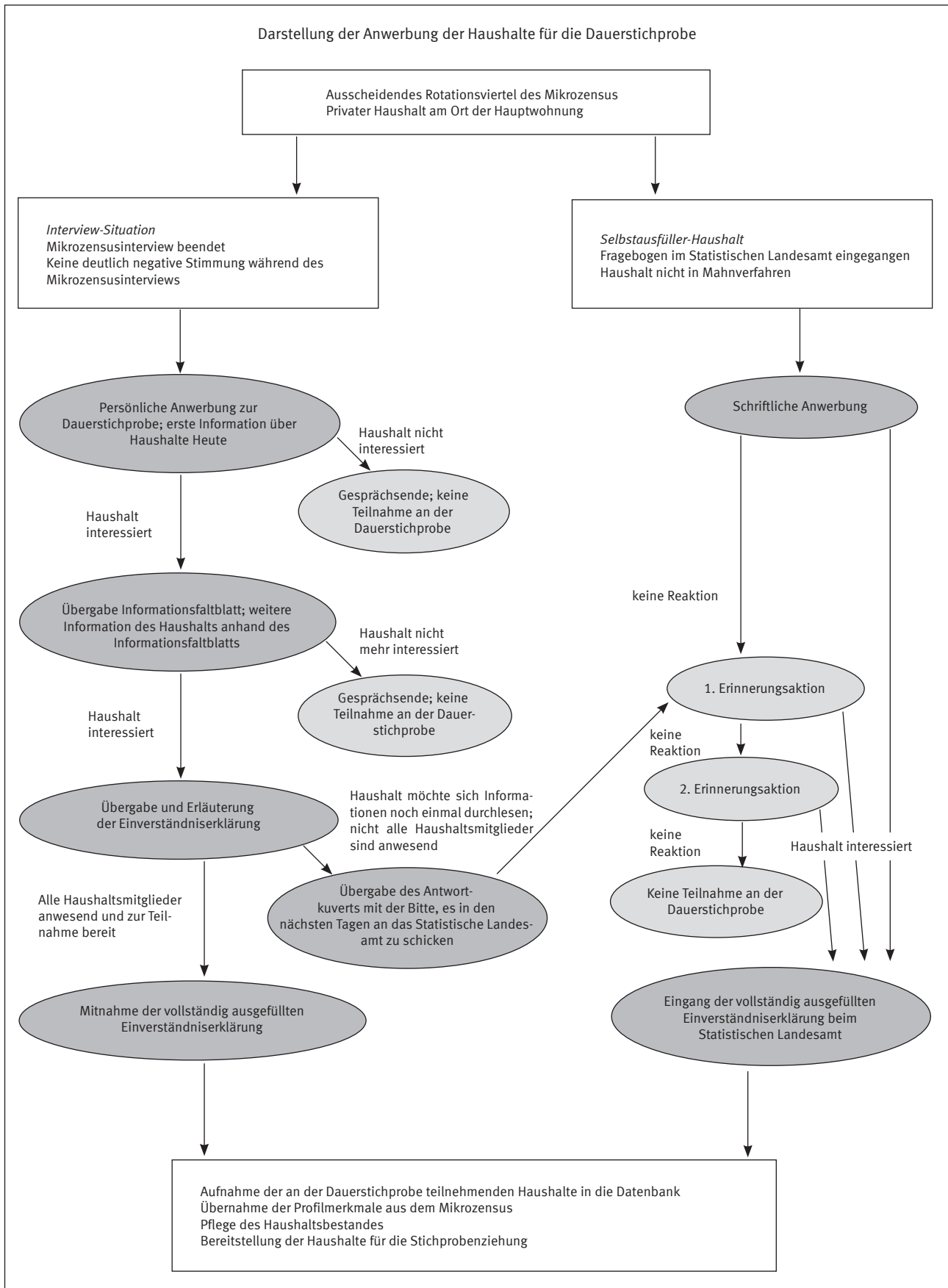
Das Konzept zum Aufbau der fünf Dauerstichprobenbausteine umfasst die Anwerbung der Haushalte für die Dauerstichprobe aus dem Mikrozensus sowie die Übernahme und Haltung zentraler Profilmkmale, die zusätzlich zu den Adressinformationen der Haushalte für die aus der Dauerstichprobe gespeisten Erhebungen zur Verfügung stehen. Darüber hinaus ist die Panelpflege von zentraler Bedeutung, deren Sinn es ist, den Kontakt zu den Haushalten über die Erhebungen hinaus zu halten, um die Panelmortalität möglichst zu minimieren. Der Baustein, welcher die Durchführung der meisten Arbeitsschritte überhaupt erst ermöglicht, ist der Aufbau einer relationalen Datenbank, welche die Haltung und Pflege der Daten regelt und eine maßgeschneiderte Verknüpfung, Filterung und Ausgabe bestimmter Datenkombinationen für die verschiedenen Arbeitsschritte und Nutzungszwecke ermöglicht. Schließlich sieht das Konzept eine umfangreiche Ausarbeitung zur Gewichtung, Hochrechnung und Varianzschätzung für Stichproben vor, die aus der Dauerstichprobe gezogen werden, um möglichst präzise Ergebnisse zu erzielen.

Das in den folgenden Abschnitten vorgestellte Gesamtkonzept bildet den Grundstein für den 2004 begonnenen Aufbau der Dauerstichprobe und regelt alle notwendigen Arbeitsschritte; es wird jedoch, besonders im Zuge der weiteren Implementierung, kontinuierlich weiterentwickelt und verbessert.

2.1 Anwerbung

Für die Dauerstichprobe werden private Haushalte an ihrem Hauptwohnsitz angeworben, die am Mikrozensus teilgenommen und ihr letztes Mikrozensusinterview abgeschlossen haben. Aufgrund des Rotationsdesigns des Mikrozensus bedeutet dies, dass jedes Jahr grundsätzlich alle Haushalte

Schaubild 3



in die Anwerbung für die Dauerstichprobe einbezogen werden, die in dem betreffenden Jahr zu dem aus dem Mikrozensus ausscheidenden Befragungsviertel gehören.

Die Anwerbung verbindet mehrere Ziele: Zum einen soll eine möglichst hohe Teilnahmequote erreicht werden. Die Erfahrungen der Pilotstudie haben gezeigt, dass hierzu die persönliche Anwerbung direkt im Anschluss an das Mikrozensusinterview am besten geeignet ist. Zugleich muss sichergestellt werden, dass die Befragten transparente Informationen darüber erhalten, was eine Teilnahme an der Dauerstichprobe für sie bedeutet und welche Daten gespeichert werden. Insbesondere ist es wichtig, dass allen Befragten klar ist, dass die Teilnahme an der Dauerstichprobe im Gegensatz zum Mikrozensus freiwillig ist und jederzeit beendet werden kann.

Auch aus diesem Grund wird darauf geachtet, dass die Haushalte erst dann angesprochen werden, wenn sie ihre letzte Mikrozensusbefragung abgeschlossen haben. Findet die Mikrozensusbefragung in einem persönlich-mündlichen Interview statt, ist dieser Zeitpunkt mit Abschluss des Interviews gegeben. Wählen die Haushalte die schriftliche Befragung des Mikrozensus (Selbstaussfüllerfragebogen), werden sie nach Eingang des Mikrozensusfragebogens im Statistischen Landesamt schriftlich angeworben.

Die überwiegende Zahl der Haushalte wird persönlich direkt im Anschluss an das letzte Mikrozensusinterview angeworben (siehe Schaubild 3). Dabei sprechen die Interviewerinnen und Interviewer des Mikrozensus die Haushalte direkt nach Beendigung des Mikrozensusinterviews auf eine Teilnahme an der Dauerstichprobe an. Bei der Anwerbung wird der Haushalt über die Dauerstichprobe informiert. Steht der Haushalt der Dauerstichprobe nicht deutlich ablehnend gegenüber, erhält er zusätzlich schriftliches Informationsmaterial in Form eines ansprechenden Faltblatts sowie eine Einverständniserklärung.

In die Dauerstichprobe werden Haushalte nur dann aufgenommen, wenn alle Haushaltsmitglieder teilnahmebereit sind, das heißt wenn alle volljährigen Haushaltsmitglieder die Teilnahmeerklärung unterschrieben haben. Dies hat methodische Gründe. Nur für vollständige Haushalte ist eine adäquate Schätzung der Teilnahmewahrscheinlichkeit möglich (siehe Abschnitt 2.5.1) und nur für vollständige Haushalte kann anschließend die Gewichtung durchgeführt werden.

Für minderjährige Haushaltsmitglieder ist die Unterschrift eines Erziehungsberechtigten ausreichend. Des Weiteren erklären sich die Haushalte mit der Unterschrift auf der Einverständniserklärung zur Übernahme einiger Informationen aus dem letzten Mikrozensusinterview (siehe Abschnitt 2.2) in die Datenbank bereit. Interessierte Haushalte haben die Möglichkeit, die Einverständniserklärung direkt zu unterschreiben und der Interviewerin oder dem Interviewer mitzugeben.

Die Interviewerinnen und Interviewer des Mikrozensus werden im Rahmen der Interviewerschulung gezielt auf die Anwerbung der Haushalte vorbereitet. Schulungsinhalte sind Ziele und Funktionsweise der Dauerstichprobe, die Argumentation gegenüber den Haushalten bei inhaltlichen Fragen zur Dauerstichprobe und zu datenschutzrechtlichen Aspekten sowie die Einweisung in die Softwareanwendung, die die Anwerbung zur Dauerstichprobe unterstützt. Gegebenenfalls wird auf Strategien zur Anwerbung schwer erreichbarer Haushalte eingegangen. Zusätzlich zur Schulung erhalten Interviewerinnen und Interviewer einen kurzen Leitfaden zur Anwerbung, welcher die Schritte der Anwerbung für die Dauerstichprobe einzeln zusammenfasst und eine Kontaktperson für Rückfragen angibt.

Haushalte, die zum Zeitpunkt der Anwerbung noch unentschlossen sind, erhalten zusätzlich zur Einverständniserklärung und zum Informationsfaltblatt einen frankierten Rückumschlag, mit dem sie bei einer positiven Entscheidung für die Teilnahme an der Dauerstichprobe die Einverständniserklärung an das jeweilige Statistische Landesamt senden können. Die noch unentschlossenen Haushalte werden etwa vier Wochen und idealerweise noch einmal acht Wochen nach dem Interview schriftlich erinnert.

Haushalten, die schriftlich am Mikrozensus teilgenommen haben, wird nach Eingang des Mikrozensusfragebogens im Statistischen Landesamt das Faltblatt und eine Einverständniserklärung zusammen mit einem Anschreiben und einem frankierten Rückumschlag zugesendet. Im Anschreiben werden die Haushalte über Idee und Ziele der Dauerstichprobe informiert und es wird um ihre Teilnahme gebeten. Senden die Haushalte die Einverständniserklärung nicht zurück, werden sie ebenfalls etwa vier Wochen nach Versand der Unterlagen und idealerweise noch einmal acht Wochen danach schriftlich erinnert.

Nach Abschluss der Anwerbung erhalten die teilnehmenden Haushalte eine Teilnahmebestätigung. Darin wird den Haushalten für ihre Teilnahme gedankt, sie erhalten nochmals Kontaktinformationen, die sie bei Fragen nutzen können, und sie werden gegebenenfalls über die nächste anstehende Erhebung informiert.

2.2 Profilvermerkmale

Von den an der Dauerstichprobe teilnehmenden Haushalten werden über die Kontaktinformationen (Adresse, Telefonnummer, gegebenenfalls E-Mail-Adresse) hinaus vor allem sozio-demografische Profilvermerkmale gespeichert, die einmal jährlich aktualisiert werden (siehe die Übersicht). Die Speicherung der Profilvermerkmale bringt für die Nutzung der Dauerstichprobe eine Reihe erheblicher Vorteile.

Die Profilvermerkmale stehen zur Schichtung der aus der Dauerstichprobe gezogenen Stichproben zur Verfügung. Durch eine geeignete Schichtung kann die Präzision der Schätzer einer Erhebung erheblich verbessert werden.⁴⁾ Jedoch ist es nur möglich, Stichproben effektiv zu schichten, wenn für die Auswahlgesamtheit relevante Schichtungsmerkmale

4) Siehe z. B. Krug, W./Nourney, M./Schmidt, J.: „Wirtschafts- und Sozialstatistik“, München, Wien 2001, S. 100 ff.

Übersicht der Profilmkmale der Dauerstichprobe

Anzahl der Personen
Haushaltsnettoeinkommen
Haushaltstyp
Soziale Stellung des Haupteinkommensbeziehers
Anzahl der Kinder unter 18 Jahren
Anzahl der Kinder von 18 bis 27 Jahren
Landwirtschaftshaushalte
Geschlecht
Geburtsmonat
Geburtsjahr
Alter
Familienstand
Staatsangehörigkeit
Stellung zum Arbeitsmarkt – Selbsteinschätzung
Erwerbstätigkeit in der Berichtswoche
Stellung im Beruf
Normalerweise geleistete Arbeitszeit je Woche
Höchster allgemeiner Schulabschluss
Höchster beruflicher Ausbildungs- oder Hochschul-/ Fachhochschulabschluss
Höhe des persönlichen Nettoeinkommens je Haushaltsmitglied
Haupteinkommensbezieher unter den teilnehmenden Personen
Beziehung zur ersten Person im Haushalt
Datum der letzten Aktualisierung der Profildaten
Gemeindekennziffer
Kennzeichnung des aktuellen Bundeslandes des Haushaltes

zur Verfügung stehen. Die in der Dauerstichprobe vorgehaltenen Profilmkmale wurden so gewählt, dass für eine Vielzahl potenzieller freiwilliger Haushalts- und Personen-erhebungen geeignete Schichtungsvariablen zur Verfügung stehen. Sie wurden insbesondere auf die Europäische Statistik über Einkommen und Lebensbedingungen (EU-SILC) abgestimmt, da für diese Erhebung mit oberster Priorität eine Auswahlgrundlage zur Verfügung gestellt werden soll.

Darüber hinaus können über die Profilmkmale aus der Dauerstichprobe Stichproben von Teilpopulationen gezogen werden, ohne dass vorher ein aufwändiges Screening durchgeführt werden muss.

Ein wesentlicher Vorteil der Dauerstichprobe besteht darin, dass mit Hilfe der Profilmkmale die Ausfallprozesse abgeschätzt werden können, welche durch die Freiwilligkeit der Teilnahme an der Dauerstichprobe und an den aus der Dauerstichprobe gespeisten Erhebungen entstehen. Über vergleichende Analysen mit den Ergebnissen des Mikrozensus erhält man eine Abschätzung der Teilnahmewahrscheinlichkeit bestimmter Schichten (siehe Abschnitt 2.5) und des auftretenden Nonresponse-Bias (Verzerrung durch Ausfälle⁵⁾). Mit deren Hilfe kann der Nonresponse-Bias bei der Gewichtung effektiv ausgeglichen werden. Folglich wurden die Profilmkmale auch so gewählt, dass ausreichende Informationen für Schätzungen der Teilnahme- und Bleibewahrscheinlichkeiten zur Verfügung stehen.

Nach Anwerbung der Haushalte aus dem Mikrozensus werden die Profilmkmale für die teilnehmenden Haushalte zunächst aus dem letzten Mikrozensusinterview in die Dauerstichprobendatenbank übernommen. Damit sie für die aus der Dauerstichprobe gespeisten Erhebungen möglichst aktuell vorliegen, werden sie einmal pro Jahr aktualisiert, das heißt erneut von den Haushalten erfragt. Um die Belastung der Haushalte möglichst gering zu halten, findet

die Aktualisierung der Profilmkmale prinzipiell im Zusammenhang mit einer Erhebung statt, an der der Haushalt teilnimmt, da ein Großteil der Profilmkmale mit den soziodemografischen Variablen übereinstimmt, die für freiwillige Haushaltserhebungen ohnehin erfragt werden. Nur wenn ein Haushalt in einem Jahr an keiner Erhebung teilnimmt, werden die Profilmkmale mit Hilfe von kurzen Haushalts- und Personen-Aktualisierungsfragebogen (separate Profildatenaktualisierung) abgefragt.

Die Übernahme der Profilmkmale aus dem Mikrozensus sowie ihre Aktualisierung im Rahmen von Erhebungen, die aus der Dauerstichprobe gespeist werden, führt zu einer erheblichen Reduzierung der Belastung für die Befragten. Darüber hinaus ist es für die Abschätzung des Nonresponse-Bias und die Schätzung von Teilnahme- und Bleibewahrscheinlichkeiten unabdingbar, dass die Profilmkmale kompatibel zu den Variablen des Mikrozensus sind.

Allerdings bedeutet dieses Konzept, dass die Profilmkmale und deren Ausprägungen sowohl an den Mikrozensus als auch an die freiwilligen Haushaltserhebungen, aus denen sie aktualisiert werden, angepasst werden müssen. Da derzeit nicht alle sozio-ökonomischen Variablen in den verschiedenen Haushaltserhebungen der amtlichen Statistik vollständig harmonisiert sind, kann es hierbei teilweise zu Problemen kommen. Die Variablen werden mit nicht immer vollständig übereinstimmenden Fragetexten und Antwortausprägungen erhoben und müssen dadurch auf verschiedene Arten typisiert werden. Weiterhin können die Aktualisierungszeitpunkte für die Profilmkmale verschiedener Dauerstichproben-Haushalte innerhalb eines Jahres unterschiedlich sein, da sich die aus der Dauerstichprobe gespeisten Erhebungen zeitlich über das Jahr verteilen. Diese Einschränkungen rechtfertigen es vor dem Hintergrund der Befragtenbelastung und des zusätzlichen Aufwandes jedoch nicht, die Profilmkmale für alle Haushalte separat zu aktualisieren. Die Vorgehensweise ist auch deshalb vertretbar, da die Profilmkmale im Wesentlichen die Funktion von Hilfsvariablen hinsichtlich der oben genannten Ziele übernehmen und nicht den Ansprüchen von Zielvariablen zur Ermittlung von Schätzern genügen müssen. Künftig sollten jedoch die wesentlichen sozio-ökonomischen Variablen im System der Haushaltserhebungen möglichst vollständig harmonisiert werden, was über die Nutzung für die Dauerstichprobe hinaus eine Reihe von Vorteilen mit sich bringen würde.

2.3 Panelpflege

Die Haushalte in der Dauerstichprobe werden nicht für eine einzelne Erhebung, sondern für eine auf einen längeren Zeitraum angelegte Kooperation angeworben. Es ist daher wichtig, dass die Teilnahme an der Dauerstichprobe für die Haushalte auch über einen längeren Zeitraum als eine interessante und sinnvolle Aufgabe wahrgenommen wird. Insbesondere muss verhindert werden, dass durch das Ausscheiden bestimmter Haushalte (sog. Panelmortalität) die Verzerrung der Ergebnisse (Bias) ein zu großes Ausmaß

⁵⁾ Siehe hierzu Biemer, P. P./Lyberg, L. E.: "Introduction to Survey Quality", Hoboken 2003, S. 80 ff.

annimmt. Um mobile Haushalte nicht zu verlieren, ist zum Beispiel eine laufende Pflege des Adressbestandes notwendig. Das Konzept zur Panelpflege in der Dauerstichprobe versucht diesen unterschiedlichen Aspekten Rechnung zu tragen.

Alle Haushalte, die an der Dauerstichprobe teilnehmen, werden auf unbestimmte Zeit weiterverfolgt.⁶⁾ Von der Weiterverfolgung werden alle Haushalte ausgeschlossen, die ihren Austritt aus der Dauerstichprobe erklären, die bei drei aufeinander folgenden Erhebungen (hierbei wird die separate Profildatenaktualisierung als Erhebung mitgezählt) die Teilnahme verweigern, ihren Hauptwohnsitz dauerhaft in einen Anstaltshaushalt oder ins Ausland verlegen oder die aufgrund eines nicht mehr nachvollziehbaren Adresswechsels nicht mehr weiterverfolgt werden können.

Um die Haushalte über wichtige Neuerungen der Dauerstichprobe zu informieren und einen stets aktuellen Adressbestand zu gewährleisten, werden alle Haushalte mindestens zweimal jährlich angeschrieben, und zwar zusätzlich zu den Erhebungen, an denen sie teilnehmen. Zu Weihnachten wird eine speziell für die Dauerstichprobe gestaltete Weihnachtskarte versendet. Im Sommer des jeweiligen Jahres wird ein Newsletter der Dauerstichprobe an alle Haushalte geschickt, in dem über die Dauerstichprobe und die aus der Dauerstichprobe gespeisten Erhebungen berichtet wird.

Umzüge von Haushalten werden in der Datenbank zusammen mit dem Datum erfasst. Ziehen einzelne Personen aus einem Haushalt aus, bilden diese zunächst einen sogenannten Splithaushalt. Über die Haushaltsidentifikationsnummer ist ersichtlich, aus welchem Ursprungshaushalt die Splithaushalte stammen, aber sie sind dennoch deutlich als eigenständige Haushalte gekennzeichnet. Die Anzahl der Umzüge eines Haushalts wird generell als Mobilitätskennzeichen registriert. Darüber hinaus wurden Regelungen für den Fall getroffen, dass zwei verschiedene Haushalte, die beide an der Dauerstichprobe teilnehmen, zusammenziehen, sodass auch dies aus der Datenbank nachvollziehbar ist.

Da die Datenbanken der Dauerstichprobe zwischen den Statistischen Landesämtern nicht verbunden sind, werden beim Umzug von Haushalten in ein anderes Bundesland die sie kennzeichnenden Datensätze aus der Anwendung des Statistischen Landesamtes des alten Wohnsitzes ausgelesen, an das entsprechende Statistische Landesamt des neuen Wohnsitzes übergeben und dort in die Datenbank importiert. Werden Haushalte nach einem Umzug nicht mehr erreicht (d. h. die Haushalte teilen keine aktuelle Kontaktadresse mit), werden sie als ausgetretene Haushalte in der Datenbank gekennzeichnet und für jegliche Stichprobenbeziehungen gesperrt.

Ebenfalls weiterverfolgt werden Haushalte oder Personen, die lediglich temporär in einen Anstaltshaushalt (z. B. Krankenhaus, Bundeswehr, Justizvollzugsanstalt) bzw. ins Ausland umziehen. Die Zeit der Unterbrechung wird in der Datenbank erfasst und der Haushalt wird durch ein Kontakt-

datum gekennzeichnet, an dem er wieder für Erhebungen zur Verfügung steht. Kurz nach Ablauf dieses Kontaktdatums wird den Haushalten ein kurzes Schreiben gesendet.

Haushaltsmitglieder, die zum Zeitpunkt der Anwerbung minderjährig waren, müssen nach Erreichen der Volljährigkeit die Einverständniserklärung ebenfalls unterzeichnen. Sollte das Einverständnis von einer betreffenden Person nicht gegeben werden, wird für die betreffende Person ein Austritt verzeichnet.

Während bei der Anwerbung grundsätzlich auf die Vollständigkeit der Haushalte bei einer Teilnahme geachtet wird (siehe Abschnitt 2.1), wird im Verlauf der Teilnahme akzeptiert, dass einzelne Haushaltsmitglieder aus der Dauerstichprobe austreten, ohne dass der Haushalt als solcher entlassen wird. Das Ausscheiden einzelner Haushaltsmitglieder kann über die Schätzung von Bleibewahrscheinlichkeiten sinnvoll erfasst werden.

Zuzüge oder Geburten in einem bestehenden Haushalt werden generell in der Datenbank vermerkt. Die betreffenden neuen Personen werden in der Datenbank zu dem bereits bestehenden Haushalt als teilnehmende Personen eingetragen, wenn sie jünger als 18 Jahre alt sind oder wenn sie älter als 18 Jahre sind und eine ihnen zugesendete Einverständniserklärung unterschrieben haben.

Austritte, Todesfälle oder die permanente Verlagerung des Hauptwohnsitzes in eine Anstalt oder ins Ausland werden in der Datenbank verzeichnet. Die entsprechenden Einheiten werden bei Stichprobenbeziehungen automatisch nicht mehr berücksichtigt und die entsprechenden Kontakt- und Identifikationsdaten (Namen, Adressen) werden automatisch gelöscht. Die Profilmerekmale werden jedoch erst dann gelöscht, wenn die Auswertungen zur Panelmortalität und Haushaltsgewichtung abgeschlossen sind. Da die Haushalte zur Speicherung der Profildaten ihr Einverständnis gegeben hatten und zu den Daten keine Kontakt- oder Namensidentifikationen mehr vorliegen, ist diese Vorgehensweise rechtlich unbedenklich.

Nimmt ein Haushalt an drei aufeinander folgenden Erhebungen (hierbei werden auch die kurzen Haushalts- und Personen-Aktualisierungsfragebogen mitgezählt, die die Haushalte erhalten, wenn sie in einem Jahr an keiner Erhebung teilgenommen haben) nicht teil, wird er aus der Dauerstichprobe entlassen, um keinen Pflegeaufwand für Haushalte betreiben zu müssen, die zwar nicht aktiv ausgetreten sind, dennoch aber kein Interesse an den Erhebungen zeigen. Eine Entlassung bewirkt in der Datenbank dieselben Sperrungen wie Austritte.

2.4 Datenhaltung

2.4.1 Grundlagen der Dauerstichprobendatenbank

Für den Aufbau und die Durchführung der Dauerstichprobe war die Schaffung einer Datenbank unabdingbar, die die Datenhaltung, Aktualisierung sowie die Panelpflege ermög-

⁶⁾ Ein aktives „Entlassen“ aller Haushalte einer Anwerbungswelle ist nur dann erforderlich, wenn der durch die Panelmortalität entstehende Bias zu groß wird. Hierüber können Aussagen erst nach dem Vorliegen einiger Erfahrungen getroffen werden. Daneben muss der Umfang der Dauerstichprobe in angemessenem Verhältnis zu Zahl und Stichprobengröße der auf ihrer Basis durchgeführten Erhebungen stehen.

lichen und gleichzeitig die Anwerbung der Haushalte unterstützen sollte. Die Arbeiten zur Programmierung wurden über den Arbeitskreis Informationstechnik der statistischen Ämter (AKIT) im Jahr 2003 an das Statistische Landesamt Baden-Württemberg vergeben.

Mit den Vorarbeiten zur Programmierung der Dauerstichprobe wurde Anfang 2004 begonnen. Es flossen sowohl die Erfahrungen, die im Rahmen der Pilotstudie gemacht wurden, als auch die Anforderungen an den bundesweiten Aufbau der Dauerstichprobe in die Überlegungen zur Realisierung ein.

Eine der wichtigsten Vorgaben des AKIT war die Begrenzung der Kosten: Die Anwendung musste unter dem kostenlosen MySQL-Server und unter der Microsoft Access-Runtimerevision lauffähig sein.⁷⁾ Die Open Source Datenbank MySQL wird frei im Sinne der „Gnu General Public Licence“ verwendet⁸⁾, das heißt es ist jedem Nutzer erlaubt, die Datenbank zu kopieren und zu vervielfältigen, solange sie nicht verändert wird.

Aufgrund des föderalen Systems der amtlichen Statistik sollte der Dauerstichprobendatenbestand getrennt in den Statistischen Landesämtern geführt werden, ohne direkte Verbindung zwischen den Datenbanken der Statistischen Landesämter. Eine zentrale Datenhaltung sowie ein Transfer personenbezogener Daten für Auswertungen oder Typisierungen waren zum Zeitpunkt der Umsetzung im Jahr 2004 nicht gewünscht.

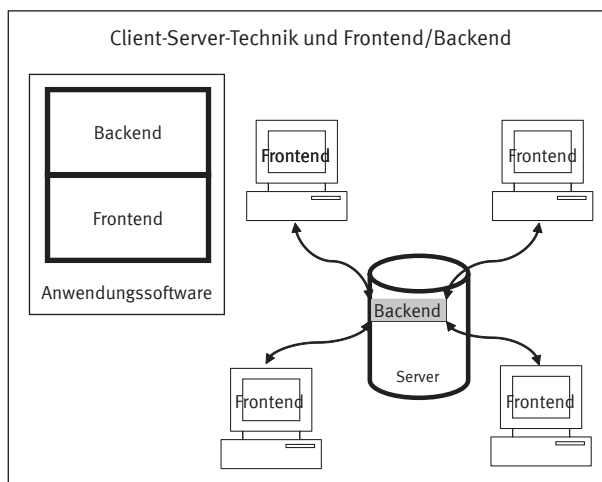
Für die Datenübertragung zwischen den Statistischen Landesämtern, die zum Beispiel bei Umzügen von Haushalten zwischen verschiedenen Bundesländern notwendig ist, wurde – da hierbei sensible Daten transferiert werden – das sichere Verfahren des Statistischen Bundesamtes DVE (DatenVersandEmpfang) gewählt. Für Großrechner war dieses Verfahren in allen Ländern bereits etabliert, für PCs wird der Datentransfer zwischen den Statistischen Landesämtern derzeit eingerichtet.

2.4.2 Realisierung der Programmierung

Die Dauerstichprobendatenbank wurde unter der Voraussetzung entwickelt, dass das Betriebssystem Microsoft Windows XP standardmäßig in allen statistischen Ämtern eingesetzt wird.⁹⁾

Zwischen MS-Access als Frontend (Access 2002-Dateiformat) und dem MySQL-Server (ab Version 4.0.17) als Backend wird über die Open Database Connectivity-Schnittstelle Version 3.51.06 kommuniziert. Als Frontend und Backend werden bestimmte Softwarekomponenten von Datenbankwendungen bezeichnet. Unter Frontend werden alle Daten und Funktionalitäten subsumiert, die die Nutzer (Clients)

Schaubild 4



sehen, als Backend werden die eigentlichen Daten bezeichnet (siehe Schaubild 4).

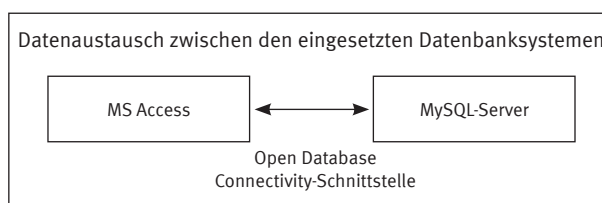
Auf dem Frontend (Client oder auch Arbeitsplatz-PC) werden die Komponenten abgelegt, die sich mit der Weiterentwicklung laufend ändern; so sind hier die eigentlichen Programme installiert.

Die Komponenten auf dem Backend (Server) ändern sich von der Struktur her normalerweise so gut wie nicht, sodass hier die Datenbankstrukturen und damit die Daten abgelegt werden.

Durch diese Einteilung können die einzelnen Nutzer parallel auf den Datenbestand zugreifen, was zum Beispiel auch eine gleichzeitige Dateneingabe ermöglicht.

Bei Open Database Connectivity (ODBC) handelt es sich um einen Standard für Datenbankzugriffe der Firma Microsoft. Dieser ist notwendig, da MS-Access nicht direkt auf die MySQL-Tabellen zugreifen kann, erst die Open Database Connectivity-Schnittstelle ermöglicht dies. Der auf dem Client-Rechner eingerichtete Treiber „MySQL ODBC 3.51 Driver DSN“ ist für das entsprechende Datenbankformat programmiert und somit in der Lage, auf die MySQL-Datenbank zuzugreifen (siehe Schaubild 5).

Schaubild 5



Für die Programmierung der Anwendung wurde die Programmiersprache Visual Basic for Applications (VBA) gewählt.

7) MySQL ist ein Relationales Datenbank Management System (RDBMS) der schwedischen Firma MySQLAB. MySQL gilt als die populärste Open-Source-Datenbank der Welt. Die MySQL AB ist der alleinige Inhaber des Quellcodes für den MySQL-Server. Somit ist MySQL kein typisches Open-Source-Projekt. Da diese Software von der MySQL AB unterstützt wird, steht aber ein einzigartiges Kosten- und Dienstleistungsmodell zur Verfügung, das eine besondere Kombination der durch Open Source gebotenen Unabhängigkeit (z. B. Anbieterunabhängigkeit) und zuverlässiger Software mit Unterstützung bietet.

8) Siehe hierzu <http://www.gnu.org/licenses/gpl.html> (Stand: 8. Mai 2006).

9) Da Windows XP erst ab Mitte 2004 in allen statistischen Ämtern zur Verfügung stand, wurden in der Einführungsphase die Programme parallel auch für Windows NT entwickelt. Sie waren damit auch in den statistischen Ämtern lauffähig, die die Umstellung des Betriebssystems noch nicht vollzogen hatten.

Aufgrund der umfangreichen und komplexen fachlichen Vorgaben wurde ein Ausbaustufenkonzept mit einzelnen Meilensteinen gewählt, das heißt Funktionalitäten und Tabellen wurden zu den Zeitpunkten bereitgestellt, zu denen sie aktuell benötigt wurden. Zum Zeitpunkt dieser Veröffentlichung ist die elfte Ausbaustufe ausgeliefert. Mit jeder weiteren Ausbaustufe geht eine permanente Weiterentwicklung – auch hinter der Nutzeroberfläche – einher.

2.4.3 Aufbau der Datenbank

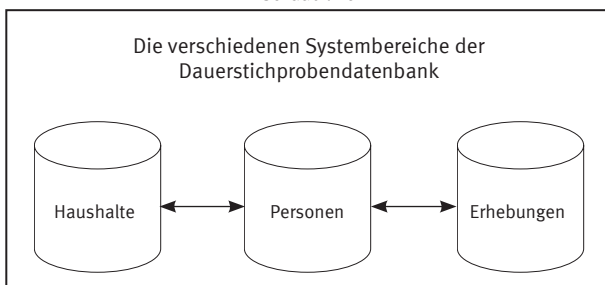
Um den Anforderungen der verschiedenen Aufgaben der Dauerstichprobe gerecht zu werden, besteht die Dauerstichprobendatenbank aus drei Bereichen (siehe Schaubild 6):

- Systembereich Haushalte
- Systembereich Personen
- Systembereich Erhebungen

Innerhalb dieser Systembereiche existieren momentan 15 MySQL-Tabellen und 90 Access-Tabellen.

Diese Gliederung in drei Systembereiche wurde gewählt, um die Beziehungen zwischen verschiedenen Daten und Datenanforderungen effizient zu gestalten und einen nutzerfreundlichen Austausch zu gewährleisten.

Schaubild 6



Der Systembereich „Haushalte“ enthält alle Kenngrößen, die zur Anwerbung der Haushalte und zur Panelpflege notwendig sind. Hierzu zählen Informationen, wie der Haushalt geworben wurde, ob ihm schriftliches Anwerbungsmaterial zugesandt werden muss, ob der Haushalt während der Anwerbung in Erinnerungsaktionen einbezogen wurde sowie zur postalischen und telefonischen Erreichbarkeit des Haushalts. Darüber hinaus werden in diesem Bereich die haushaltsbezogenen Profilm Merkmale vorgehalten.

Der Systembereich „Personen“ enthält in erster Linie die personenbezogenen Profilm Merkmale. Darüber hinaus wird hier der Aktualitätsstand der Profilm Merkmale festgehalten, um sie rechtzeitig zu aktualisieren, falls dies nicht im Zusammenhang mit einer Erhebung erfolgt. Ebenso wird vom Systembereich Person eine Verbindung zum Systembereich Erhebung geschaffen, um zu registrieren, welche Einheiten zu welchen Erhebungen herangezogen und kontaktiert wurden.

Da die Personen immer einem Haushalt zugeordnet werden können, ist eine Generalisierung der Informationen für den

Haushalt, zum Beispiel hinsichtlich der Aktualität der Profilm Merkmale oder der Erhebungsteilnahme, möglich.

Im Systembereich „Erhebungen“ werden alle Charakteristika der einzelnen Erhebungen geführt. Dazu gehören der Beginn und das Ende der Feldphase, die Belastungsintensität und Informationen zu den Haushalten bzw. Personen, die für die jeweilige Erhebung und deren Fortschreibung gezogen wurden.

Ebenso wird hier die Teilnahme oder Nichtteilnahme der für eine Erhebung ausgewählten Personen verzeichnet.

2.4.4 Nutzeroberflächen und Funktionalitäten

Über die Nutzeroberflächen werden alle Datenimporte und -exporte sowie alle für Anwerbung und Pflege notwendigen Funktionalitäten bereitgestellt. Um die Nutzung und den Zugang zur Dauerstichprobe auf verschiedenen Ebenen jedoch verwalten zu können, werden verschiedene Nutzerkategorien festgelegt.

Über diese Nutzerverwaltung wird geregelt, wer auf welche Funktionalitäten der Dauerstichprobendatenbank zugreifen kann. Das Konzept umfasst ein vierstufiges Nutzersystem mit folgenden Gruppen:

- Dauerstichprobenadministrator
- Fach-Administrator
- Bearbeiter
- Erfasser

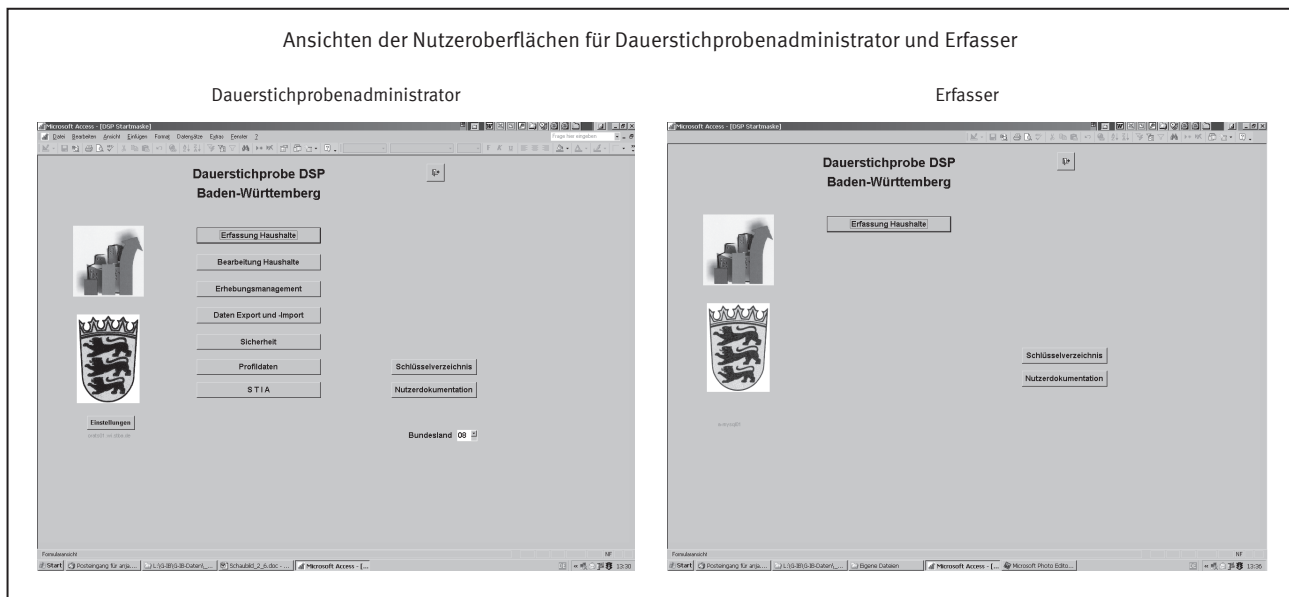
Für die Gruppe der Erfasser ist die Ersterfassung der Haushalte und Personen möglich. Die Bearbeiter können zusätzlich die eigentlichen Personen- und Haushaltsdaten korrigieren und bearbeiten, zum Beispiel eine Adressänderung oder Austritte eintragen. Die Fachadministratoren haben uneingeschränkten Zugang zu allen Funktionalitäten der Datenbank. Die Dauerstichprobenadministratoren können zusätzlich den einzelnen Nutzern der vorgenannten Gruppen die Zugangsrechte zuweisen oder entziehen. Die Authentifikation der Nutzergruppen erfolgt über Passwörter, welche den Sicherheitsbestimmungen der Statistischen Landesämter genügen. Jede Nutzergruppe erhält eine individuelle Sicht auf die Module der Dauerstichprobe.

In Schaubild 7 sind anhand der Benutzeroberfläche des Administrators alle Funktionalitätengruppen der Dauerstichprobe dargestellt. Über diese Funktionalitäten können alle Arbeitsschritte, die in den Abschnitten 2.1 bis 2.3 beschrieben sind, durchgeführt werden.

In „Erfassung Haushalte“ ist die Erfassung neuer Haushalte und deren Personen entweder über die Einverständniserklärung manuell oder medienbruchfrei direkt aus der Blaiseanwendung für die Anwerbung möglich.

Unter dem Punkt „Bearbeitung Haushalte“ werden alle Änderungen der Haushalts- und Personendaten, aber auch Umzüge innerhalb eines Bundeslandes oder in ein anderes Bundesland bzw. ins Ausland, temporär oder auf

Schaubild 7



Dauer, bearbeitet. Hierunter fallen auch Haushaltssplittings, Haushaltsfusionen und interne Abgleiche, die der Datenkonsistenz dienen (wie der der Postleitzahl mit der Gemeindekennziffer).

„Erhebungsmanagement“ ist die Verwaltung der Erhebungen an sich sowie die Informationseingabe über Teilnahme oder Nichtteilnahme von Haushalten bzw. Personen an Erhebungen manuell oder maschinell. Darüber hinaus besteht eine Abfragemöglichkeit dieser Informationen.

„Daten-Export und -Import“ umfasst den Import aller bisher oder in Zukunft separat erfassten Haushalts- oder Personen-daten über eine definierte Schnittstelle mit Prüfroutine in die Dauerstichprobe. Insbesondere galt dies für die bereits aus der Pilotstudie bereitgestellten Daten. Besondere Recherchen nach Haushalten, Personen oder Erhebungen sowie diversen umfangreichen Kriterien ermöglichen den Export ganz spezieller Datengruppen. Die Serienbriefherstellung für diverse Anschreiben ist möglich.

Unter „Sicherheit“ wird eine eigene Datenbanksicherung der Dauerstichprobe durch den Nutzer zu jedem beliebigen Zeitpunkt bereitgestellt, und zwar zusätzlich zur und unabhängig von der Sicherung auf Systemseite. Die Datenbankwiederherstellung mit diesen sensiblen Daten ist zu einem beliebigen, vom Dauerstichprobenadministrator bestimmten Zeitpunkt mit dieser Funktionalität möglich.

Beim Punkt „Profildaten“ erfolgt die Übernahme der Profilmerkmale aus dem Mikrozensus nach Abschluss der jeweiligen Anwerbungswelle. Unterschieden werden die jeweiligen Datenstrukturen des Mikrozensus der einzelnen Erhebungszeiträume. Unmittelbar danach wird die Verbindung zum Mikrozensus getrennt. Eine spätere Profildatenaktualisierung speziell aus EU-SILC, aber auch aus anderen Erhebungen, sowie eine separate Profildatenaktualisierung ohne direkten Zusammenhang mit einer bestimmten Erhebung erfolgen in dieser Funktionalitätengruppe. Die Aus-

gabe der Profilmerkmale (ohne Namens- und Kontaktinformationen) ist möglich.

Unter dem letzten Punkt der Nutzeroberfläche (STIA) wird die Stichprobenziehung von Haushalten bzw. Personen für die Teilnahme an Erhebungen abgehandelt.

Die einzelnen Funktionalitäten sind weitgehend selbsterklärend gestaltet; zusätzlich kann jedoch auf eine umfangreiche Nutzerdokumentation online zugegriffen werden.

2.4.5 Installation

Die initiale Installation der Dauerstichprobendatenbank erfolgt über ein Setup. Da die Voraussetzungen in den einzelnen Statistischen Landesämtern unterschiedlich waren, wurden verschiedene Setups bereitgestellt: für Windows XP mit Access-Voll- bzw. Runtime-Version, für Windows NT mit Access-Voll- bzw. Runtime-Version. Die dabei verwendete EasyPHP-Software installierte alle für die Dauerstichprobe benötigten Komponenten (einschl. Apache-Server, PHP, MySQL, PHPmyAdmin), sofern nicht in einzelnen Ämtern ein eigenes Konzept zur Administration von MySQL realisiert war.

Bei der Installation der weiteren Ausbaustufen wird entweder nur die Frontendseite aktualisiert oder es wird bei einer Erweiterung der MySQL-Struktur zusätzlich das Backend gesichert, die MySQL-Datenbank mit Hilfe einer separaten Datei (SQL-Script) neu angelegt und die Sicherung zurückgeladen. Structured Query Language (SQL) ist eine Definitions- und Abfragesprache für relationale Datenbanken.

Im Rahmen der amtlichen Statistik wurde für die Dauerstichprobendatenbank erstmals Microsoft Access in Verbindung mit MySQL angewandt. Aus diesem Grund wird bei der Programmierung großer Wert auf Service im Sinne von Kundendienst gegenüber den Anwendern in den Statistischen Landesämtern gelegt. Das Dauerstichproben-Entwicklungsteam im Statistischen Landesamt Baden-Württemberg sieht

sich als technischer Dienstleister für die Anwendung der Dauerstichprobe und nicht nur als reiner Entwickler der Datenbank und deren Programme.

In der Einführungsphase wurde intensive Hilfestellung für alle Statistischen Landesämter – neben der Parallelentwicklung der Programme auch für Windows NT – geleistet, damit der Einstieg in Datenhaltung und -pflege der Dauerstichprobendatenbank erleichtert wurde.

Hierzu gehörten auch Einführungsschulungen, die nach Auslieferung der ersten Ausbaustufen zusammen mit dem Statistischen Bundesamt in drei Bundesländern durchgeführt wurden. Darüber hinaus werden Anregungen, Hinweise und Wünsche an die Programmierung während der gesamten Entwicklung mit dem bekannten Nutzerkreis abgestimmt.

2.5 Hochrechnung

Allgemeines Ziel der Hochrechnung ist es, mit Hilfe geeigneter Schätzfunktionen von den Stichprobenwerten auf die Parameter der Grundgesamtheit zu schließen. Die Hochrechnung für Erhebungen aus der Dauerstichprobe erfordert vorab Untersuchungen und Berechnungen von Wahrscheinlichkeiten, um den komplexen Aufbau der Dauerstichprobe entsprechend zu berücksichtigen. Speziell müssen die Haushalte und Personen bezüglich ihrer Teilnahme an Erhebungen aus der Dauerstichprobe (Teilnahmewahrscheinlichkeiten) und bezüglich ihres Verbleibs in der Dauerstichprobe (Bleibewahrscheinlichkeiten) betrachtet werden.

2.5.1 Teilnahmewahrscheinlichkeiten

Im Folgenden soll zuerst die Teilnahmewahrscheinlichkeit an der Dauerstichprobe genauer untersucht werden.

Wie bereits in Abschnitt 2.1 erwähnt, werden jährlich alle Haushalte aus dem Mikrozensus, die im betreffenden Jahr zu dem Viertel der Haushalte zählen, das aus dem Mikrozensus heraus rotiert, für die Dauerstichprobe angesprochen. Die Wahrscheinlichkeit, im betreffenden Jahr für die Dauerstichprobe angeworben zu werden, beträgt somit ein Viertel der Wahrscheinlichkeit, im betreffenden Jahr im Mikrozensus vertreten zu sein. Da im Mikrozensus jährlich 1% der Gesamtbevölkerung Deutschlands befragt werden, beträgt diese Wahrscheinlichkeit also 0,25%.

Zwar besteht im Mikrozensus Auskunftspflicht, doch beträgt der Nonresponse jährlich etwa 3%. Dieser verteilt sich wie bei den meisten anderen Erhebungen nicht gleichmäßig über alle Schichten, sondern ist für die Haushaltsgruppen unterschiedlich. Aus diesem Grund wird jährlich zusätzlich zur Hochrechnung des Mikrozensus ein so genannter „Kompensationsfaktor der Ausfälle“ berechnet, der neben der Anzahl der Haushaltsmitglieder deren Geschlecht, Alter und Staatsangehörigkeit berücksichtigt. Dieser Kompensationsfaktor wird für alle zur Teilnahme an der Dauerstichprobe bereiten Personen aus dem Mikrozensus übernommen und mit dem Faktor von 0,25% multipliziert. Das Produkt stellt die Wahrscheinlichkeit dar, für die Dauerstichprobe angesprochen zu werden.

Die Teilnahmebereitschaft an der Dauerstichprobe ist, wie bei allen freiwilligen Erhebungen, nicht bei allen Personen und Haushalten gleich. So nehmen beispielsweise Beamte häufiger an freiwilligen Erhebungen teil als Selbstständige. Bei Personen mit einem niedrigen Einkommen oder einem hohen Einkommen ist die Teilnahmebereitschaft generell geringer und Einpersonenhaushalte nehmen seltener als Haushalte von Paaren mit Kind(ern) teil. Auch bei den Anwerbungen für die Dauerstichprobe wurde dies beobachtet. Da für teilnahmebereite Personen und Haushalte neben dem Bundesland, dem Alter, dem Geschlecht, dem Familienstand und dem Haushaltstyp weitere sozio-ökonomische Variablen, wie beispielsweise das Haushaltsnettoeinkommen, die Stellung im Beruf, der Schulabschluss und der berufliche Abschluss (siehe die Übersicht auf S. 457), aus dem Mikrozensus übernommen werden, können über den Abgleich der im Mikrozensus und der in der Dauerstichprobe vertretenen Personen und Haushalte mittels logistischer Regression die Teilnahmewahrscheinlichkeiten an der Dauerstichprobe für alle teilnahmebereiten Personen und Haushalte berechnet werden. Um eine Übereinstimmung der aus dem Mikrozensus betrachteten Personen- und Haushaltsinformationen zu gewährleisten, ist es wichtig, dass Informationen für alle Personen eines Haushalts vorliegen. Für die Berechnungen gilt:

$$\text{Teilnahme: } Y = 1 \quad \text{Prob}(Y = 1) = F(\beta^T x)$$

$$\text{Nicht-Teilnahme: } Y = 0 \quad \text{Prob}(Y = 0) = 1 - F(\beta^T x)$$

Im Logit-Modell stellt F die logistische Verteilung dar. Die Teilnahmewahrscheinlichkeit wird daher folgendermaßen berechnet:

$$\text{Logit-Modell: } \text{Prob}(Y = 1) = \frac{e^{\beta^T x}}{1 + e^{\beta^T x}} = \Lambda(\beta^T x).$$

Als erklärende Variablen x dienen hier beispielsweise das Bundesland, das Alter, der Haushaltstyp, das Haushaltsnettoeinkommen. Zu beachten ist, dass Variablen mit einem möglichst hohen Erklärungsgrad gewählt werden und die Anzahl der Variablen im Hinblick auf mögliche Auswirkungen auf die Varianz begrenzt wird. Untersuchungen zum geeignetsten Modell werden derzeit noch im Rahmen eines Kooperationsprojektes mit den Universitäten Tübingen und Trier unter der Leitung von Prof. Ralf Münnich durchgeführt. Da die erste bundesweite Anwerbungswelle erst 2004 erfolgte, liegen derzeit nur Ergebnisse von zwei bundesweiten Anwerbungen vor. Die erklärenden Variablen zur Berechnung der Teilnahmewahrscheinlichkeit werden daher jährlich auf ihre Erklärungskraft untersucht und möglicherweise noch durch andere ersetzt. Ziel ist es, ein Modell zu finden, das die Wahrscheinlichkeit, an der Dauerstichprobe teilzunehmen, möglichst optimal erklärt.

2.5.2 Bleibewahrscheinlichkeiten

Wie in Abschnitt 2.2 beschrieben, werden jährlich die Profildaten aller Personen und Haushalte, die in der Dauerstichprobe vertreten sind, entweder durch eine Erhebung oder eine separate Profildatenaktualisierung aktualisiert. Dies

stellt zum einen sicher, dass für künftige Erhebungen die Profildaten für alle teilnahmebereiten und für alle an der Erhebung nicht teilnehmenden Personen bekannt sind und diese somit für die Stichprobenziehung und die Berechnung der Teilnahmewahrscheinlichkeiten an der Erhebung mittels logistischer Regression genutzt werden können. Des Weiteren können diese Angaben verwendet werden, um am Ende jedes Jahres die Wahrscheinlichkeit, noch weiter teilnahmebereit zu sein, zu ermitteln. Analog zur Berechnung der Teilnahmewahrscheinlichkeit wird die Bleibewahrscheinlichkeit mittels logistischer Regression berechnet, wobei auch hier die Profilm Merkmale als erklärende Variablen dienen. Die Berechnung erfolgt ebenfalls auf Personen- und Haushaltsebene.

Da im Laufe eines Jahres auch nur ein Teil eines Haushalts seine Teilnahme beenden kann, wird zusätzlich ein weiteres Merkmal eingeführt, das kennzeichnet, ob der vollständige Haushalt weiter teilnahmebereit ist oder ein Teil des Haushaltes seine Bereitschaft zurückgezogen hat. Der Austritt eines Teils der Haushaltsmitglieder aus der Dauerstichprobe kann sich negativ auf die Bereitschaft der übrigen Haushaltsmitglieder auswirken und wird daher bei weiteren Berechnungen berücksichtigt.

Auch die räumliche Mobilität spielt beim Verbleib der Person/des Haushalts in der Dauerstichprobe eine Rolle. In Zukunft soll daher die Anzahl der Umzüge bei den Berechnungen der Bleibewahrscheinlichkeit mit berücksichtigt werden.

Die berechneten Teilnahme- und Bleibewahrscheinlichkeiten werden stets dem aktuellen Dauerstichprobendatensatz zugeführt. Für jede aus der Dauerstichprobe gezogene Erhebung stehen diese Angaben somit zur Verfügung und können bei der Hochrechnung verwendet werden.

2.5.3 Hochrechnung

Bei der Hochrechnung wird von der Stichprobe, die aus der Dauerstichprobe gezogen wurde, auf die Grundgesamtheit geschlossen. Am Beispiel von EU-SILC 2005 soll das Hochrechnungskonzept vorgestellt werden. Zur Hochrechnung werden in der Regel Gewichtungsfaktoren wie folgt berechnet:

$$w_k = \frac{g_k}{\pi_k \cdot \hat{\theta}_k}$$

mit

w_k : Hochrechnungsfaktor für Person k /Haushalt k

g_k : Korrekturfaktor für Person k /Haushalt k

π_k : Wahrscheinlichkeit für Person k /Haushalt k , für die Erhebung EU-SILC 2005 ausgewählt zu werden

$\hat{\theta}_k$: geschätzter Faktor der Teilnahmewahrscheinlichkeit an der Erhebung EU-SILC 2005 für Person k /Haushalt k

Die Möglichkeit, an der Erhebung EU-SILC 2005 teilzunehmen, wird von der Wahrscheinlichkeit, in der Dauerstichprobe zu sein, bestimmt.

Die Wahrscheinlichkeit π_k besteht somit aus mehreren Faktoren. Sie ist das Produkt aus:

- der Wahrscheinlichkeit, im herausrotierenden Rotationsviertel des Mikrozensus zu sein (0,25% mal Kompensationsfaktor für Ausfälle),
- der Wahrscheinlichkeit, an der Dauerstichprobe teilzunehmen (mittels logistischer Regression geschätzte Teilnahmewahrscheinlichkeit),
- der Wahrscheinlichkeit, zum Erhebungszeitpunkt in der Dauerstichprobe vertreten zu sein (Produkt der jährlichen mittels logistischer Regression geschätzten Bleibewahrscheinlichkeiten),
- der Wahrscheinlichkeit einer Person k /eines Haushalts k , die/der in der Dauerstichprobe vertreten ist, für die Erhebung EU-SILC 2005 ausgewählt zu werden. Diese Wahrscheinlichkeit π_k für eine Person k /für einen Haushalt k ergibt sich über den Auswahlsatz der Erhebung [sie wird für jede Schicht über den Quotienten der Anzahl der Personen in der Stichprobe und der Grundgesamtheit (Dauerstichprobe) berechnet bzw. über die Anzahl der Haushalte in der Stichprobe und der Grundgesamtheit (Dauerstichprobe)].

Der Faktor $\hat{\theta}_k$ ist die Wahrscheinlichkeit einer ausgewählten Person k /eines ausgewählten Haushalts k , an der Erhebung EU-SILC 2005 teilzunehmen, das heißt die mittels logistischer Regression geschätzte Teilnahmewahrscheinlichkeit der Person k /des Haushaltes k an dieser Erhebung.

Erst das Produkt $\pi_k \cdot \theta_k$ erklärt dann die Wahrscheinlichkeit insgesamt für die Person k /den Haushalt k , in der Erhebung EU-SILC 2005 vertreten zu sein.

Der Korrekturfaktor g_k wird durch die Anpassung der Stichprobe an Eckwerte der Bevölkerung berechnet, um so mögliche durch Nonresponse entstehende Verzerrungen auszugleichen. Die Bestimmung des Korrekturfaktors g_k erfolgt durch eine verallgemeinerte Regressionsschätzung und ergibt sich aus folgender Gleichung:

$$g_k = 1 + (t_x - \hat{t}_{x,HT})' \left(\sum_{k=1}^n \frac{x_k x_k'}{\pi_k \hat{\theta}_k} \right)^{-1} x_k,$$

wobei

x_k der Vektor aller Ausprägungen der Hilfsmerkmale für Person k /Haushalt k und

t_x der Vektor der Totalwerte der Hilfsmerkmale

ist.

Der Horwitz-Thompson-Schätzer der Totalwerte wird folgendermaßen berechnet:

$$\hat{t}_{x,HT} = \sum_{k=1}^n \frac{x_k}{\pi_k \hat{\theta}_k}$$

Somit ergibt sich als lineare Schätzfunktion für einen Totalwert t_y der GREG-Schätzer (Generalized Regression Estimator) folgendermaßen:

$$\begin{aligned} \hat{t}_y &= \hat{t}_{y,HT} + \hat{B}'(t_x - \hat{t}_{x,HT}) \\ &= \sum_{k=1}^n (1 + (t_x - \hat{t}_{x,HT})' (\sum_{k=1}^n d_k x_k x_k')^{-1} x_k) d_k y_k \\ &= \sum_{k=1}^n w_k y_k \end{aligned}$$

Bisher wurden für die Hochrechnung der EU-SILC-Daten von 2005 die Teilnahmewahrscheinlichkeiten an der Dauerstichprobe für alle 2004 angeworbenen Personen berechnet sowie deren Bleibewahrscheinlichkeiten für das Jahr 2004. Da für die Erhebung EU-SILC 2005 nur Daten von 2004 verwendet werden konnten und die Anwerbungsergebnisse für 2005 zum Zeitpunkt der Veröffentlichung dieses Artikels noch nicht vollständig vorlagen, konnte die Berechnung der Teilnahmewahrscheinlichkeiten an der Dauerstichprobe 2005 noch nicht durchgeführt werden. Um $\hat{\theta}_k$ vollständig berechnen zu können, wurden für alle Personen, die für EU-SILC 2005 aus dem Dauerstichproben-Datensatz von 2004 angesprochen wurden, die Teilnahmewahrscheinlichkeiten an der Erhebung 2005 berechnet. Somit kann das Produkt $\pi_k \cdot \hat{\theta}_k$ vollständig berechnet werden und dient bei der Hochrechnung der EU-SILC-Daten 2005 als Eingangsgewicht für die Kalibrierung.

3 Nutzungsmöglichkeiten der Dauerstichprobe

Die Nutzungsmöglichkeiten der Dauerstichprobe sind vielfältig und erstrecken sich grundsätzlich auf alle Themenbereiche, die im Rahmen der amtlichen Statistik mit freiwilligen Befragungen von Haushalten oder Personen abgedeckt werden können. Die wesentliche Grundvoraussetzung für die Speisung einer Erhebung aus der Dauerstichprobe ist die Freiwilligkeit der Erhebung. Die Möglichkeit der Stichprobenziehung aus der Dauerstichprobe hängt darüber hinaus von den Anforderungen der Erhebung bezüglich Stichprobenumfang und Schichtungsvorgaben ab. Das heißt die Realisation einer Stichprobe aus der Dauerstichprobe muss im Vorfeld mit der in der Dauerstichprobe befindlichen Auswahlgrundlage verglichen werden, um potenzielle Probleme für die Genauigkeit der gewonnenen Schätzer ausschließen zu können. So eignet sich die Dauerstichprobe in der Regel nicht für Erhebungen bei sehr kleinen Teilpopulationen (z. B. allein erziehende Väter), da diese nicht in hinreichendem Umfang angeworben werden können.

Einen allgemeinen Überblick über mögliche Stichprobenumfänge und Grundgesamtheiten gibt eine Analyse der Anwerbungsergebnisse. Für das Anwerbungsjahr 2004 liegen bereits ausführliche Auswertungen vor. Darüber hinaus können erste Aussagen für das Anwerbungsjahr 2005

getroffen werden, für welches derzeit allerdings noch nicht alle Anwerbungsinformationen vorliegen.

Ein erstes Bild der Bandbreite von Nutzungsmöglichkeiten der Dauerstichprobe bieten die Erhebungen, welche bereits aus der Dauerstichprobe gespeist werden.

Schließlich bestehen auch eine Reihe von Bedingungen für die Nutzung der Dauerstichprobe, die sich aus momentan bestehenden Limitationen, aber auch durch die Rahmenbedingungen der amtlichen Statistik ergeben. Diese Bedingungen werden abschließend in diesem Kapitel zusammengefasst.

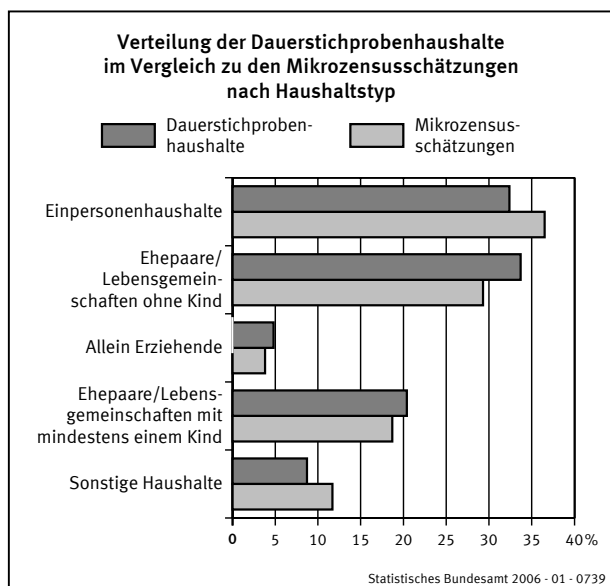
3.1 Anwerbungsergebnisse der Dauerstichprobe

Im Jahr 2004 wurde erstmals bundesweit für die Dauerstichprobe angeworben, wobei über 7 700 Haushalte ihre Teilnahmebereitschaft erklärten. Gemessen an allen im Jahr 2004 aus dem Mikrozensus ausgeschiedenen Haushalten entspricht dies einer bundesweiten Teilnahmequote von 9,4%. Die Anwerbungserfolge in den einzelnen Bundesländern waren trotz einheitlicher Methodik der Anwerbung jedoch sehr unterschiedlich; die Teilnahmequoten in den einzelnen Ländern liegen zwischen 5,5 und 15,1%. Im Jahr 2005 wurden in einigen Ländern bereits Teilnahmequoten von etwa 20% erzielt. Dies zeigt, dass das Potenzial für die Teilnahme an der Dauerstichprobe bislang nicht ausgeschöpft wird. Ziel der Anwerbung in den nächsten Jahren ist daher, bundesweit Teilnahmequoten von deutlich mehr als 10% zu erzielen. Ein Hauptgrund für die differierenden Teilnahmequoten wird in der unterschiedlichen Vorbereitung der Interviewerinnen und Interviewer auf die Dauerstichprobenanwerbung gesehen. Ein wichtiges Ziel beim weiteren Aufbau der Dauerstichprobe in den nächsten Jahren ist daher, gemeinsam die besten Vorgehensweisen im Hinblick auf Umfang und Inhalte der Schulung zu identifizieren und bundesweit einheitlich anzuwerben, um gute Teilnahmequoten zu erzielen.

In der Dauerstichprobe werden die Haushalte aus den verschiedenen Anwerbungsjahren kumuliert. Auch aus der Pilotstudie zur Dauerstichprobe konnten bereits Haushalte übernommen werden, sodass Ende 2004 bereits über 10 500 Haushalte in der Dauerstichprobenbank verzeichnet waren. Für das Anwerbungsjahr 2005 wird von ähnlichen oder besseren Teilnahmequoten ausgegangen; derzeit befinden sich ungefähr 20 000 Haushalte in der Dauerstichprobe. Bei einem angestrebten jährlichen Zuwachs von 10 000 Haushalten sollten Anfang des Jahres 2008 knapp 40 000 Haushalte zur Stichprobenziehung aus der Dauerstichprobe zur Verfügung stehen.

Ein weiteres Kriterium für die Möglichkeiten zur Nutzung der Dauerstichprobe ist die zu erwartende Verzerrung für die jeweils interessierenden Erhebungsmerkmale. Erhebungen aus der Dauerstichprobe können nur dann zuverlässige Schätzer über die interessierende Grundgesamtheit liefern, wenn die Verzerrung durch Antwortausfälle (Nonresponse-Bias) in der Auswahlgrundlage Dauerstichprobe nicht zu groß ist und effektiv korrigiert werden kann.

Schaubild 8



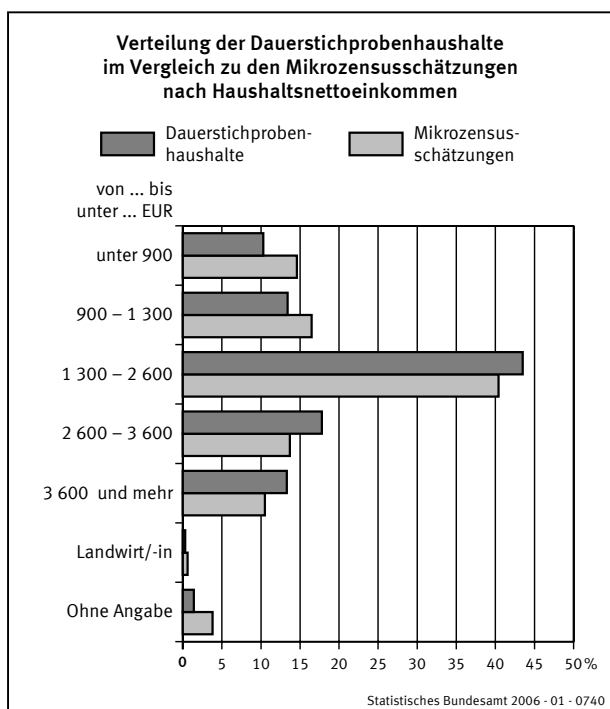
Erste Einschätzungen über die Verteilung innerhalb der Auswahlgrundlage erhält man über Auswertungen der Profilmerkmale. Ein einfacher Vergleich der Verteilungen dieser Variablen bei den Dauerstichprobenhaushalten und in den Schätzungen des Mikrozensus, auch Nettovalidierung¹⁰⁾ genannt, liefert erste Anhaltspunkte über das Ausmaß der

durch die Freiwilligkeit der Teilnahme entstehenden Verzerrungen. Beispielhaft dargestellt werden hier die Verteilungen von Haushaltstyp (siehe Schaubild 8) und Haushaltsnettoeinkommen (siehe Schaubild 9) der Haushalte in der Dauerstichprobe (Stand: Anfang 2005).

Generell zeigen sich im Vergleich zum Mikrozensus auch in der Dauerstichprobe die bei freiwilligen Erhebungen üblicherweise vorkommenden Verzerrungen. Diese fallen jedoch – trotz der niedrigen Teilnahmequoten – insgesamt gering aus und können durch geeignete Schätzverfahren korrigiert werden (siehe auch Abschnitt 2.5.1 Teilnahmewahrscheinlichkeiten). Erste Ergebnisse von Studien zur Gewichtung sowie zur Varianzschätzung zeigen, dass die Dauerstichprobe als Auswahlgrundlage den Genauigkeitsanforderungen der meisten freiwilligen Haushaltserhebungen der amtlichen Statistik genügen wird.¹¹⁾

Trotz dieser guten Ausgangssituation für allgemeine Bevölkerungsbefragungen sollten für eine Reihe schwer anzuwerbender Bevölkerungsgruppen Verbesserungen der Anwerbungsergebnisse angestrebt werden. Hierzu zählen insbesondere Haushalte mit sehr niedrigen Einkommen sowie Personen ohne einen oder nur mit niedrigem Bildungsabschluss. Darüber hinaus sind selbstständige Personen und Personen, die als berufliche Stellung Arbeiterin/Arbeiter angeben, in der Dauerstichprobe nur unterdurchschnittlich vertreten. Die Dauerstichprobe enthält darüber hinaus zu wenige Personen, die nicht die deutsche Staatsbürgerschaft besitzen. Altersverteilungen zeigen ferner, dass besondere Anstrengungen erforderlich sind, um junge Menschen zwischen 20 und 29 Jahren verstärkt für eine Teilnahme zu gewinnen.

Schaubild 9



Neben der Erhöhung der Teilnahmequoten ist ein weiteres Ziel, die daraus entstehenden Verzerrungen so weit als möglich zu minimieren. Hierbei wird berücksichtigt, dass über höhere Teilnahmequoten nicht generell eine Verringerung des Nonresponse-Bias erzielt werden kann, sondern dass Anwerbungsmethoden immer beide Ziele gleichermaßen im Auge behalten müssen. Ein effektives Monitoring der Anwerbungserfolge hinsichtlich ihrer Auswirkung auf den Nonresponse-Bias findet über die Abschätzung des relativen Nonresponse-Bias mit einem vereinfachten Modell nach Biemer und Lyberg¹²⁾ statt.

3.2 Erhebungen aus der Dauerstichprobe

Die Dauerstichprobe wurde primär als Auswahlgrundlage für die Umsetzung der neuen Europäischen Statistik über Einkommen und Lebensbedingungen (EU-SILC) aufgebaut. EU-SILC verfolgt das Ziel, europaweit statistische Informationen über Einkommen, Armut und soziale Ausgrenzung zu gewinnen, die zum Beispiel als Grundlage für die Entwicklung von Strategien zur Armutsbekämpfung und zur Überprüfung von deren Wirksamkeit wichtig sind.¹³⁾

10) Siehe Schneekloth, U./Leven, I.: „Woran bemisst sich eine ‚gute‘ allgemeine Bevölkerungsumfrage? Analysen zu Ausmaß, Bedeutung und zu den Hintergründen von Nonresponse in zufallsbasierten Stichprobenerhebungen am Beispiel des ALLBUS“ in ZUMA-Nachrichten 53, Mannheim 2003, S. 16 ff.

11) Siehe Münnich, R./Huergo, L./Magg, K./Ohly, D.: „Konzeption und Test von Varianzschätzung für Erhebungen auf Basis der Dauerstichprobe befragungsbereiter Haushalte“, Abschlussbericht, Tübingen 2005, S. 92.

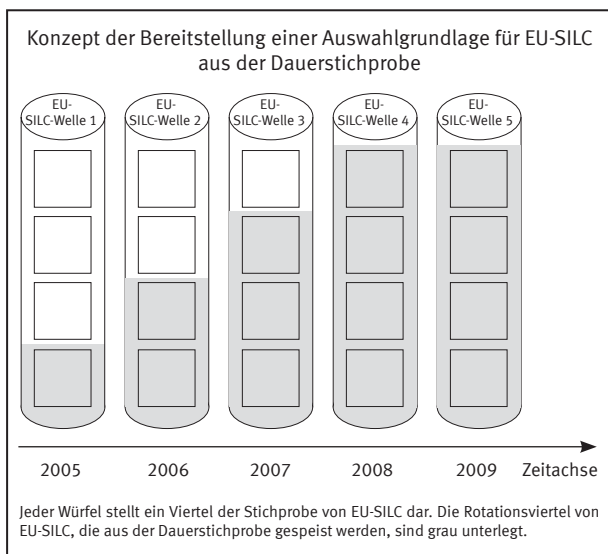
12) Siehe Fußnote 5.

13) Näheres zu EU-SILC siehe Körner, T./Meyer, I./Minkel, H./Timm, U.: „LEBEN IN EUROPA – Die neue Statistik über Einkommen und Lebensbedingungen“ in WiSta 11/2005, S. 1137 ff.

Der Nettostichprobenumfang von EU-SILC beträgt etwa 14 100 Haushalte. Die Erhebung wurde erstmals im Jahr 2005 durchgeführt. Da die Dauerstichprobe allerdings erst ab dem Jahr 2004 aufgebaut wurde, konnte nicht der gesamte Stichprobenumfang aus der Dauerstichprobe gespeist werden. Aus diesem Grund wurde im Jahr 2005 nur ein Teil der Erhebungseinheiten für EU-SILC über die Stichprobenziehung aus der Dauerstichprobe gewonnen.

EU-SILC wird als Vierjahresrotationspanel realisiert. Dies bedeutet, dass die Haushalte vier Jahre in EU-SILC verbleiben und jedes Jahr ein Viertel der Stichprobenhaushalte ausscheidet. An diesem Konzept orientiert sich die Nutzung der Dauerstichprobe für die Stichprobenziehung: Die EU-SILC-Stichprobe wird von 2005 bis 2007 in zwei Teile unterteilt: in einen Teil, der aus der Dauerstichprobe gespeist wird, und in einen Teil, der aus anderen Quellen als Quotenstichprobe realisiert wird. Im Jahr 2005 wurden ein Viertel der Haushalte für EU-SILC aus der Dauerstichprobe gewonnen und drei Viertel aus anderen Quellen. Von letzteren Haushalten scheidet jedes Jahr bis einschließlich 2007 ein Viertel aus EU-SILC aus und wird durch ein „frisches“ Viertel aus der Dauerstichprobe ersetzt, sodass erstmals im Jahr 2008 EU-SILC vollständig aus der Dauerstichprobe gespeist wird (siehe Schaubild 10).

Schaubild 10



Die Belastung der Befragten bei EU-SILC ist relativ hoch. Die Erhebung wird schriftlich durchgeführt und die Haushalte erhalten je einen Haushaltsfragebogen sowie für jede Person im Haushalt einen Personenfragebogen, in denen detailliert Einkommenskomponenten erfragt werden. Für den Haushaltsfragebogen und jeden Personenfragebogen werden durchschnittlich jeweils etwa 30 Minuten zum Ausfüllen benötigt. Gemessen an dieser Belastung ist die Teilnahmequote für die aus der Dauerstichprobe teilnehmenden Haushalte hoch: 65% aller aus der Dauerstichprobe angeschriebenen Haushalte nahm 2005 an EU-SILC teil. Um einen

Nettostichprobenumfang von ungefähr 4 000 Haushalten zu erzielen, mussten also ungefähr 6 200 Haushalte angeschrieben werden.

Eine weitere Erhebung, die aus der Dauerstichprobe gespeist wird, ist die Europäische Umfrage zur Nutzung von Informations- und Kommunikationstechnologien (IKT). Für die Piloterhebungen zu dieser Statistik wurde bereits die Pilotstudie zur Dauerstichprobe genutzt. Mit Hilfe der IKT-Erhebung werden zum Beispiel Aussagen über die Internetnutzung in Deutschland, den Umgang mit dem Internet in verschiedenen Teilen der Bevölkerung, über die Art, wie diese Bevölkerungsschichten die für den Internetzugang notwendigen Kenntnisse und Fertigkeiten erwerben, und darüber, in welchem Umfang das Internet von Privatpersonen für Onlineeinkäufe genutzt wird, gewonnen. Einen weiteren Schwerpunkt bildet die Nutzung des Internets für die Inanspruchnahme von Dienstleistungen der öffentlichen Verwaltung (E-Government¹⁴). Da die IKT-Erhebung als Quotenstichprobe realisiert wird und den Statistischen Landesämtern freigestellt ist, welche Quellen sie nutzen, wird die Erhebung nur zum Teil aus der Dauerstichprobe gespeist. Die Belastung der Befragten ist im Vergleich zu EU-SILC wesentlich geringer und entsprechend werden für den aus der Dauerstichprobe gespeisten Anteil noch höhere Teilnahmequoten als bei EU-SILC erzielt. Die Teilnahmequote der Dauerstichprobenhaushalte für die IKT-Erhebung 2005 lag bei ungefähr 71%. Dies bestätigt die guten Ergebnisse für die Teilnahmequoten bei den IKT-Erhebungen, die aus der Pilotstudie zur Dauerstichprobe gespeist wurden, von über 80% (2002 und 2003 jeweils 81%).

Die erste personenbezogene Erhebung, die aus der Dauerstichprobe gespeist wird, ist die Erhebung demografischer Grunddaten zur Geburtenentwicklung. Diese Erhebung richtet sich ausschließlich an Frauen im Alter zwischen 16 und 75 Jahren. Ziel der Erhebung ist es, zuverlässige Angaben zu wichtigen Indikatoren der Geburtenentwicklung, wie etwa zur Kinderzahl je Frau, zum Alter der Mütter bei Geburt ihrer Kinder sowie zur Kinderlosigkeit zu gewinnen. Für die Erhebung erhalten die Frauen einen kurzen vierseitigen Personenfragebogen, der inhaltlich jeweils auf die Altersgruppen zwischen 16 und 49 Jahren bzw. 50 und 75 Jahren zugeschnitten ist. Die Feldphase der Erhebung ist auf Herbst 2006 festgelegt. Erste Ergebnisse der Erhebung werden Anfang 2007 erwartet.

3.3 Nutzungsbedingungen für künftige Erhebungen

Die Nutzungsmöglichkeiten der Dauerstichprobe sind sowohl auf Bundes- als auch auf Landesebene vielfältig. Der Einsatz kann in den nächsten Jahren noch ausgebaut werden. Grundsätzlich können Haushalts- oder Personenerhebungen der amtlichen Statistik aus der Dauerstichprobe gespeist werden, wenn sie freiwillig sind und ihr Stichprobenumfang den Rahmen der Dauerstichprobe nicht übersteigt. Damit bietet sich die Nutzung der Dauerstichprobe

14) Siehe Timm, U./Kahle, I.: „E-Government und andere Zwecke der Internetnutzung“ in WiSta 7/2005, S. 717 ff; ein weiterer Beitrag in dieser Zeitschrift zur PC- und Internetnutzung in privaten Haushalten und zur Aneignung von E-Skills ist in Vorbereitung.

insbesondere für die Durchführung von Erhebungen für besondere Zwecke nach § 7 BStatG an, die zur Erfüllung eines kurzfristig auftretenden Datenbedarfs oberster Bundesbehörden genutzt werden. Für die Nutzung der Dauerstichprobe als Auswahlgrundlage müssen jedoch eine Reihe von Rahmenbedingungen beachtet werden:

Ein wesentliches Entscheidungskriterium ist die Belastung der Haushalte durch die Gesamtheit der Erhebungen, die in einem bestimmten Jahr aus der Dauerstichprobe gezogen werden. Erste Erfahrungen aus der Pilotstudie zeigen, dass die Panelmortalität von Haushalten, die zu keiner oder nur einer Erhebung angeschrieben wurden, höher ist als die von Haushalten, die an mehreren Erhebungen teilgenommen haben.¹⁵⁾ Dies deutet darauf hin, dass diejenigen Haushalte, die sich für die Teilnahme an der Dauerstichprobe entschlossen haben, ein „reichhaltiges Angebot“ an Erhebungen mit verschiedenen Themen begrüßen und eher enttäuscht sind, wenn sie nach ihrer Teilnahmeerklärung nur sporadisch auf ihre Teilnahme an Erhebungen angesprochen werden. Detaillierte Analysen über die Auswirkungen der Anzahl an Erhebungen je Haushalt auf den Verbleib in der Dauerstichprobe können allerdings erst durchgeführt werden, wenn die Dauerstichprobe einige Jahre besteht.

Nach Schnell¹⁶⁾ hängt die Belastung von Befragten wesentlich von der Dauer einer Erhebung, vom Aufwand bei der Beantwortung einzelner Fragen, vom Ausmaß, in dem Fragen als unangenehm empfunden werden, und von der Befragungshäufigkeit ab. Für Zwecke des Monitorings der Belastung der befragten Dauerstichprobenhaushalte werden daher die Teilnahmen an Erhebungen zusammen mit einem Belastungsfaktor, der den unterschiedlichen Charakter der Erhebungen berücksichtigt, in der Dauerstichprobendatenbank verzeichnet. Unterschieden werden Erhebungen mit hoher, mittlerer und geringer Belastungsintensität. Der Belastungsfaktor eines Haushalts je Jahr setzt sich aus der Häufigkeit der Teilnahmen dieses Haushalts an verschiedenen Erhebungen und der Belastungsintensität dieser Erhebungen zusammen. Wenn eine nicht zu vernachlässigende Anzahl von Haushalten in einem Jahr einen kritischen Belastungswert erreicht hat, sollten in der Regel keine weiteren Erhebungen mehr aus der Dauerstichprobe gespeist werden. Zudem wird darauf geachtet, dass zwischen zwei Erhebungen zumindest ein bis zwei Monate liegen. Bei neuen Erhebungen wird sich der Zeitpunkt der Feldphase daher auch an den Anforderungen der Dauerstichprobe orientieren müssen.

Die Belastungsintensität der Erhebungen wird anhand der Faktoren Dauer der Befragung, Aufwand für die Beantwortung einzelner Fragen und Anzahl an Fragen, die als unangenehm empfunden werden können, eingeschätzt. Die Dauer einer Erhebung hängt zunächst von der Länge des Fragebogens und der Anzahl der Fragen ab. Hierbei sollte allerdings nicht nur die Anzahl der Fragen berücksichtigt werden, sondern auch der Schwierigkeitsgrad oder Aufwand für deren

Beantwortung. Fragen, die ohne Zuhilfenahme von Unterlagen und ohne längeres Überlegen beantwortet werden können, sind demnach weniger belastungsintensiv als Fragen, die nur durch Nachschlagen (z. B. von Gehaltsabrechnungen) oder durch längeres Nachdenken beantwortet werden können. Das Ausmaß, in dem Fragen als unangenehm empfunden werden (sensible Fragen), wird – soweit möglich – mit Hilfe qualitativer Pretests abgeschätzt.

Allerdings hängt die von den Befragten wahrgenommene Belastung nicht nur von der Dauer der Erhebung, sondern auch von der Einstellung der Befragten zur Erhebung ab. Hierbei spielen folgende Einflussgrößen eine wesentliche Rolle: das Interesse am Thema der Befragung, Ziele der Statistik und die durchführende Organisation bzw. der Auftraggeber.¹⁷⁾ Da mit der Dauerstichprobe nur Erhebungen im Rahmen der amtlichen Statistik durchgeführt werden, das heißt die durchführende Organisation eine konstante Größe ist, werden Thema und Ziel der Befragung als weitere Einflussfaktoren auf die Befragtenbelastung berücksichtigt.

Da die Häufigkeit der Befragung einzelner Dauerstichprobenhaushalte je Jahr vom Stichprobenumfang der Erhebungen relativ zum Gesamtumfang der Dauerstichprobe abhängt, ist – zumindest in den Jahren des Aufbaus der Dauerstichprobe – der Stichprobenumfang ebenfalls ein Entscheidungskriterium für die Durchführung von Erhebungen. Da der Umfang der Dauerstichprobe jedoch mit den Jahren anwächst, wird damit gerechnet, dass in einigen Jahren auch mehrere Erhebungen mit größeren Stichprobenumfängen (etwa vergleichbar mit EU-SILC) durchgeführt werden können, da sich die Belastung dann gleichmäßig auf die Dauerstichprobenhaushalte verteilt.

Für die Nutzung der Dauerstichprobe wurden bestimmte Prioritäten festgelegt: Mit oberster Priorität werden Erhebungen gespeist, die aufgrund nationaler oder europäischer rechtlicher Verpflichtungen durchgeführt werden müssen (z. B. EU-SILC). Danach werden Erhebungen auf der Grundlage von § 7 Abs. 1 BStatG (sowie vergleichbare länderspezifische oder europäische Erhebungen) und danach Erhebungen auf der Grundlage von § 7 Abs. 2 BStatG (sowie vergleichbare länderspezifische Erhebungen) bedient. Anschließend können prinzipiell auch andere Erhebungen gespeist werden.

Da Erhebungen der amtlichen Statistik grundsätzlich eine Rechtsgrundlage erfordern, können weitere Einschränkungen hinzukommen. So müssen Erhebungen für besondere Zwecke nach § 7 BStatG kurzfristig auftretenden Datenbedarf einer obersten Bundesbehörde decken oder zur Klärung wissenschaftlich-methodischer Fragestellungen auf dem Gebiet der Statistik erforderlich sein. Diese Statistiken dürfen eine Obergrenze von höchstens 20 000 Befragten nicht übersteigen und die Auskunfterteilung muss freiwillig sein. Wiederholungsbefragungen sind bis zu fünf Jahre nach der ersten Befragung zulässig. [UU](#)

15) Siehe Nimmergut, A./Meyer, I./Körner, T.: „HAUSHALTE HEUTE 2003 – Pilotstudie zur Umsetzbarkeit einer Dauerstichprobe befragungsbereiter Haushalte in der amtlichen Statistik 2001 – 2003“, Wiesbaden 2004, S. 99.

16) Siehe hierzu Schnell, R.: „Nonresponse in Bevölkerungsumfragen – Ausmaß, Entwicklung und Ursachen“, Opladen, 1997, S. 176.

17) Siehe hierzu Groves, R. M./Fowler, F. J./Couper, M. P./Lepkowski, J. M./Singer, E./Tourangeau, R.: „Survey Methodology“, Hoboken 2004.

Auszug aus Wirtschaft und Statistik

© Statistisches Bundesamt, Wiesbaden 2006

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.

Herausgeber: Statistisches Bundesamt, Wiesbaden

Schriftleitung: Johann Hahlen
Präsident des Statistischen Bundesamtes
Verantwortlich für den Inhalt:
Brigitte Reimann,
65180 Wiesbaden

- Telefon: +49 (0) 6 11/75 20 86
- E-Mail: wirtschaft-und-statistik@destatis.de

Vertriebspartner: SFG Servicecenter Fachverlage
Part of the Elsevier Group
Postfach 43 43
72774 Reutlingen
Telefon: +49 (0) 70 71/93 53 50
Telefax: +49 (0) 70 71/93 53 35
E-Mail: destatis@s-f-g.com

Erscheinungsfolge: monatlich



Allgemeine Informationen über das Statistische Bundesamt und sein Datenangebot erhalten Sie:

- im Internet: www.destatis.de

oder bei unserem Informationsservice
65180 Wiesbaden

- Telefon: +49 (0) 6 11/75 24 05
- Telefax: +49 (0) 6 11/75 33 30
- www.destatis.de/kontakt