

## Regionaldaten der amtlichen Statistik – Ein Angebot im Spannungsfeld rechtlicher Restriktionen, statistisch-methodischer Beschränkungen und dem Bedarf an regional tief gegliederten Daten

### Teil 2: Möglichkeiten und Grenzen der Regionalisierbarkeit



Von Dr. Stefan Weil

Die Nachfrage nach regionalisierten Ergebnissen der amtlichen Statistiken ist groß. Planer und Entscheidungsträger auf kommunaler und auf Landesebene sowie in Unternehmen benötigen solche Ergebnisse ebenso wie z. B. Wissenschaftler, die sich mit Regionalforschung befassen. Die amtliche Statistik verfügt über ein breites Angebot an Regionaldaten, die in Form von Tabellen, Grafiken und Karten sowie über kundenorientierte Auswertungen angeboten werden. Der Regionalisierbarkeit von Erhebungsergebnissen sind jedoch Grenzen gesetzt. In der Februarausgabe des Statistischen Monatshefts wurde in einem ersten Teil dieses Beitrags das Angebot des Statistischen Landesamtes an regionalisierten Ergebnissen dargestellt. Dieser zweite Teil befasst sich mit der grundsätzlichen Frage, welche Möglichkeiten und Grenzen bei der Regionalisierung von Ergebnissen der amtlichen Statistiken bestehen.

#### **Möglichkeiten und Grenzen der Regionalisierung von Ergebnissen amtlicher Erhebungen**

Räumliche  
Bezugnahme  
auf Ein-  
heiten der  
Verwaltungs-  
gliederung

In der Regel wird bei jeder amtlichen Erhebung ein räumlicher Bezug der erhobenen Merkmale hergestellt. Üblich sind hierbei geografische Einheiten, die sich aus der hierarchischen Verwaltungsgliederung ergeben.

NUTS 1–5

Zu diesem Zweck wird in der amtlichen Statistik die 1980 vom Europäischen Amt für Statistik (Eurostat) entwickelte Systeme

der Gebietseinheiten für die Statistik (Nomenclature des unités territoriales statistiques – NUTS) verwendet. Nach dieser Nomenklatur bilden die Gemeinden (NUTS 5) die niedrigste administrative Gliederungsebene für Erhebungen in der amtlichen Statistik. Die nächsthöhere Ebene stellen in Rheinland-Pfalz die Verbandsgemeinden (NUTS 4) dar, gefolgt von den kreisfreien Städten (gleichzeitig verbandsfreie Gemeinden) und Landkreisen (NUTS 3). Die höchste Gliederungsebene direkt unterhalb des Landes bilden in der amtlichen Statistik

**T 1** Gliederung des Landes Rheinland-Pfalz nach der Systematik der Gebietseinheiten für die Statistik (NUTS)

Ebene	Bezeichnung	Zahl der geografischen Einheiten
NUTS 1	Bundesland	1
NUTS 2	statistische Regionen <sup>1)</sup>	3
NUTS 3	kreisfreie Städte und Landkreise	36
NUTS 4 (LAU 1) <sup>2)</sup>	Verbandsgemeinden	163
NUTS 5 (LAU 2) <sup>2)</sup>	verbandsfreie und verbandsangehörige Gemeinden	2 306

1) Die „statistischen Regionen“ entsprechen den ehemaligen Regierungsbezirken Koblenz, Trier und Rheinhessen-Pfalz. Die Regierungsbezirke wurden in Rheinland-Pfalz im Jahr 2000 aufgelöst. – 2) Die Ebenen NUTS 4 bzw. NUTS 5 wurden im Zuge einer im Juli 2003 in Kraft getretenen Verordnung in LAU 1 bzw. LAU 2 („Local Area Units“) umbenannt.

die sogenannten „statistischen Regionen“ (NUTS 2). In Rheinland-Pfalz entsprechen diese den ehemaligen Regierungsbezirken, die Anfang 2000 als Verwaltungsebene aufgelöst wurden.

Tabelle 1 zeigt die NUTS-Gliederung für das Land Rheinland-Pfalz.

Auch andere Strukturräume lassen sich abbilden

Darüber hinaus lassen sich auch andere Strukturräume abbilden. Dazu zählen z. B.

- die Mikrozensus-Regionen („Anpassungsschichten“),
- die landesplanerischen Regionen („Planungsregionen“),
- die Agenturbezirke der Bundesagentur für Arbeit,
- die Kammerbezirke der Industrie- und Handelskammern bzw. der Handwerkskammern und
- die Tourismusregionen.

Während sich bestimmte Strukturräume aus den geografischen Gliederungen der Verwaltungseinheiten durchgängig ableiten – wie beispielsweise die Mikrozensus-

Regionen und die Planungsregionen –, ist dies bei anderen (z. B. den Agenturbezirken) nur eingeschränkt gegeben. So lassen sich sowohl die Mikrozensus-Regionen als auch die Planungsregion grenzscharf aus kreisfreien Städten und Landkreisen und damit auch aus den Gemeinden des Landes zusammenfassen (aggregieren). Die Agenturbezirke hingegen können lediglich durch eine von der Gemeindeebene ausgehende Aggregation abgebildet werden. Über die Zusammenfassung von Kreisen ist dies nicht möglich. Umgekehrt lassen sich die anderen angeführten Strukturräume nicht aus den Agenturbezirken aggregieren. Die regionalen Gliederungen der kommunalen Verwaltung und der Bundesagentur für Arbeit sind also nur eingeschränkt „kompatibel“.

Eine weitergehende geografische Gliederung der Erhebungsergebnisse ist in der deutschen amtlichen Statistik – im Gegensatz zu anderen europäischen Ländern – bisher in der Regel nicht möglich. Die amtliche Statistik in Deutschland strebt jedoch an, die Nutzung von georeferenzierten Daten mit einem kleinräumigen Bezug zu ermöglichen. Hierbei soll ein Konzept zur Datenhaltung, Analyse und Verbreitung von Geofachdaten zum Einsatz kommen, das sich in der amtlichen Statistik vieler europäischer Staaten bereits bewährt hat. Hierdurch sollen raumbezogene Daten z. B. über natürliche Personen, Haushalte oder Unternehmen bundesweit harmonisiert auch unterhalb der Gemeindeebene unter Beachtung der nach dem Datenschutz und der statistischen Geheimhaltung vorgegebenen Restriktionen genutzt werden können.<sup>1)</sup>

Nachweis regionalisierter Erhebungsergebnisse unterhalb der Gemeindeebene bislang noch nicht möglich

1) Vgl. ausführlich Szibalski, M.: Kleinräumige Bevölkerungs- und Wirtschaftsdaten in der amtlichen Statistik Europas. Ergebnisse einer Umfrage zur Speicherung, Analyse und Publikation. In: Wirtschaft und Statistik 2/2007, S. 137–143.

## Welche Faktoren beeinflussen die Regionalisierbarkeit der Daten der amtlichen Statistik?

Bei der Erhebung und Auswertung von Daten müssen in der amtlichen Statistik verschiedene Rahmenbedingungen beachtet werden. Hierzu gehören im Wesentlichen

- rechtliche Bestimmungen,
- die Erhebungsart,
- die Anzahl der zu erhebenden Einheiten,
- die Akzeptanz bei den Auskunftspflichtigen sowie
- die Kosten der Erhebungen.

Insbesondere die ersten drei Aspekte haben Einfluss auf das Regionalisierungspotenzial der in den Erhebungen gewonnenen Daten. Aber auch die letzten beiden sind zumindest indirekt von Relevanz, da sie in einem – zumeist kausalen – Bezug zu den anderen Aspekten stehen.

Auskunftspflicht und Datenschutz stellen eine hohe Qualität der Daten sicher

Zwei wesentliche Teilaspekte der rechtlichen Rahmenbedingungen sind die Auskunftspflicht und der Datenschutz. Beide sind im Bundesstatistikgesetz (BStatG) geregelt.

Die für den überwiegenden Teil der amtlichen Statistiken bestehende Auskunftspflicht (§ 15 BStatG) garantiert eine hohe Rücklaufquote und damit – selbst bei tiefer fachlicher und regionaler Gliederung – aussagekräftige Ergebnisse. Statistiken, bei denen keine Verpflichtung zur Auskunft besteht, weisen in der Regel hohe Antwortausfälle mit entsprechenden Auswirkungen auf die Qualität der Erhebungsergebnisse auf (siehe Textkasten Auskunftspflicht).

Das Pendant zur Auskunftspflicht bildet die statistische Geheimhaltung: Einzel-

## Auskunftspflicht

Verlässliche statistische Informationen bedürfen der Auskunftspflicht bei der Erhebung, da die bzw. der Befragte in der Regel kein Eigeninteresse an deren Preisgabe hat (Kollektivgutproblematik). Aus diesem Grunde werden Bürgerinnen und Bürger, Unternehmen und andere Institutionen gesetzlich verpflichtet, Auskünfte zu amtlichen Statistiken zu erteilen. Rechtlich beinhaltet die Auskunftspflicht, dass die bzw. der Befragte, die Auskunft vollständig, wahrheitsgemäß, unentgeltlich und fristgerecht erteilen muss und dass dies gegebenenfalls mittels staatlichen Zwangs durchsetzbar ist. Die Auskunftspflicht tangiert allerdings das Grundrecht auf informationelle Selbstbestimmung. Deshalb darf keine Erhebung der amtlichen Statistik ohne gesetzliche Ermächtigung erfolgen.

Die Notwendigkeit, vollständiger und zuverlässiger Information wächst mit in dem Maße wie aus den Erhebungen der amtlichen Statistik regionalisierte und/oder fachlich gegliederte Ergebnisse gewonnen werden sollen. Dieses Erfordernis besteht insbesondere dann, wenn keine Vollerhebungen, sondern nur Teilerhebungen durchgeführt werden können.

Der Informationswert von Erhebungsergebnissen, die auf freiwilliger Basis gewonnen werden, ist in der Regel reduziert. Dies liegt darin begründet, dass die Beteiligungsquote bei freiwilligen Erhebungen meistens geringer ist als bei einer Erhebung mit Auskunftspflicht. Untersuchungen des Statistischen Bundesamtes zeigen, dass bei Betrieben des produzierenden Gewerbes mit Ausfallquoten bis zu 50% gerechnet werden muss, wenn die Teilnahme an der Erhebung freigestellt wird. Zudem wirkt sich aus, dass sich bei hohen Nonresponse-Quoten die einzelnen Gruppen der Grundgesamtheit in der Regel nicht repräsentativ an der Erhebung beteiligen. Wenn aber manche Gruppen stärker und andere schwächer beteiligt sind, führt dies zu Verzerrungen, die die Aussagekraft der Statistiken in elementarer Weise mindern.<sup>1)</sup>

1) Vgl. Westerhoff, Horst-Dieter: Die amtliche Statistik in der demokratischen Gesellschaft. In: Wirtschaft und Statistik, H. 11/2007, S. 1130–1145.

angaben über persönliche und sachliche Verhältnisse, die für eine amtliche Statistik gemacht werden, sind gemäß § 16 BStatG geheim zu halten, soweit durch besondere Rechtsvorschrift nichts anderes bestimmt ist. Bei jeder Veröffentlichung von geheimhaltungskritischen Daten wird geprüft, ob der Schutz vor der Offenlegung von Einzelangaben garantiert ist (siehe Textkasten Geheimhaltung).<sup>2)</sup>

2) Vgl. auch Statistische Ämter des Bundes und der Länder (Hrsg.): Qualitätsstandards in der amtlichen Statistik, o. O. 2003., S. 3.

## Geheimhaltung

Ein zentraler Pfeiler der amtlichen Statistik ist die Gewährleistung der statistischen Geheimhaltung. Alle Auskunftspflichtigen müssen darauf vertrauen können, dass ihre individuellen Daten geheim bleiben. Damit ist die statistische Geheimhaltung gewissermaßen eine Gegenleistung für die Auskunftspflicht. Andererseits ist mit dem Schutz der Daten gegen ihre Offenlegung untrennbar ein Informationsverlust verbunden, der die Aussagefähigkeit der veröffentlichten Statistik (wenn auch auf kontrollierbare Weise) reduziert.

In Auswertungen und Veröffentlichungen ist stets dafür Sorge zu tragen, dass geheimhaltungskritische Daten von der Anzeige ausgeschlossen werden. Maßnahmen zur Gewährleistung der Geheimhaltung können hierbei an den aggregierten Ergebnissen einer statistischen Aufbereitung ansetzen oder sich auf das Ausgangsmaterial (die Mikrodaten) einer Statistik beziehen. Im ersten Fall wird von Tabellengeheimhaltung (auf der Grundlage sogenannter Zellsperungsverfahren oder Rundungsverfahren) im zweiten Fall von Anonymisierung des Ausgangsdatenmaterials gesprochen.<sup>1)</sup>

Im Rahmen der Tabellengeheimhaltung wird die „Unterdrückung“ dieser Tabellenfelder als „primäre Geheimhaltung“ bezeichnet. Danach sind z. B. Aggregationen von Daten mit weniger als drei Einzelfällen oder mit dominanten Einzelwerten problematisch. Neben der primären Sperrung von Tabellenfeldern müssen in der Regel zusätzliche Werte sekundär gesperrt werden, um eine Aufdeckung der primär gesperrten Zellen im Rahmen von Gegenrechnungen (z. B. einfache Differenzbildung bei Zwischensummen aufweisenden Tabellen) zu verhindern.

Mikrodaten hingegen, wie sie beispielsweise von den FDZ für die wissenschaftliche Nutzung angeboten werden („Scientific Use Files“), müssen durch gezielte Veränderungen (z. B. Weglassen, Vergrößern oder Vertauschen von Merkmalen, Ziehen von Stichproben usw.) zumindest „faktisch anonymisiert“ werden. So wird sichergestellt, dass die ausgewiesenen Einzelfälle vor „Deanonymisierung“ durch die Datennutzer geschützt sind.

Das Grundprinzip der Anonymisierung besteht darin, durch (merkmalsträger-, merkmals- oder ausprägungsbezogene) Informationsreduktion oder Informationsveränderung (Datenverände-

rung) eine Reidentifikation der Merkmalsträger weitgehend zu verhindern. Informationsreduktion erfolgt bei den merkmalsbezogenen Verfahren durch das Entfernen auffälliger Merkmalsträger („Ausreißer“), systematische Einschränkung der Grundgesamtheit (Entfernen einer Teilgesamtheit) oder durch Ziehen einer (Sub-)Stichprobe. Bei merkmalsbezogenen Verfahren werden Merkmale beseitigt (Variablenunterdrückung), ersetzt (z. B. durch Kennziffern oder Indizes) oder zusammengefasst bzw. Merkmalsausprägungen vergrößert (z. B. durch Gruppierung in Größenklassen, Rundung). Beim ausprägungsbezogenen Vorgehen werden hingegen einzelne Werte (von Ausprägungen, die in der Stichprobe sehr selten vorkommen oder gar einzigartig sind) unterdrückt.

Werden lediglich merkmalsträgerbezogene Informationen wie Name, Anschrift oder z. B. Sozialversicherungsnummer unterdrückt, liegt eine „formale Anonymisierung“ vor. Formal anonymisierte Mikrodaten werden von den FDZ lediglich im Rahmen der kontrollierten Datenfernverarbeitung bereitgestellt. „Faktische Anonymisierung“ bedeutet hingegen, dass Reidentifikationsbemühungen durch weitere Informationsreduktion so weit erschwert werden, dass sie sich – rationales Verhalten der potenziellen Nutzer vorausgesetzt – ökonomisch nicht lohnen (Reidentifikationsaufwand ist größer als der Nutzen der zusätzlich gewonnenen Information).

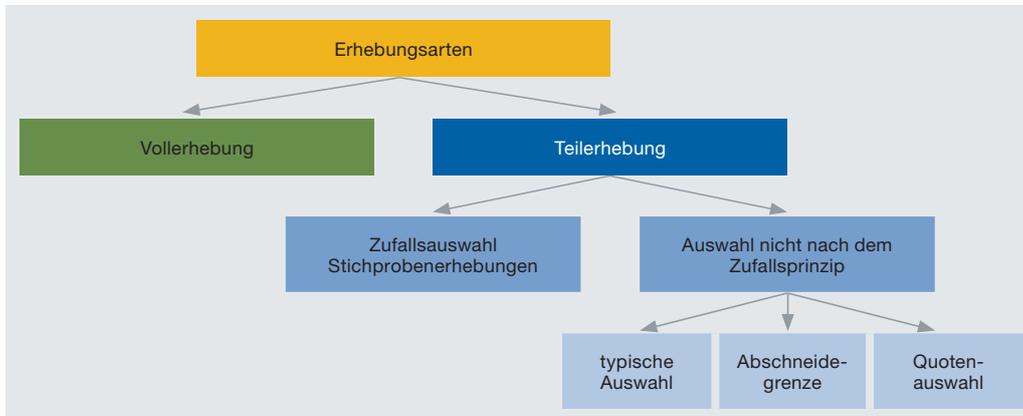
Das faktisch anonymisierte Datenmaterial verfügt häufig nur noch über ein eingeschränktes Analysepotenzial.<sup>2)</sup> Sehr stark reduziert ist das Analysepotenzial jedoch in den für einen breiteren Nutzerkreis aufbereiteten vollständig anonymisierten Datensätzen in den sogenannten „Public Use Files“.

1) Vgl. auch Heitzig, Jobst: Möglichkeiten zur statistischen Geheimhaltung bei komplexen Analysen amtlicher Mikrodaten in Forschungsdatenzentren. In: METHODEN – VERFAHREN – ENTWICKLUNGEN, Nachrichten aus dem Statistischen Bundesamt, Ausgabe 1/2005, S. 7–8.

2) Zu Anonymisierungsverfahren und ihrer Auswirkung auf das Analysepotenzial siehe ausführlich Ronning, Gerd, u. a.: Handbuch zur Anonymisierung wirtschaftsstatistischer Mikrodaten. Statistik und Wissenschaft, Bd. 4, hrsg. v. Statistischen Bundesamt. Wiesbaden 2005, und Rosemann, Martin: Auswirkungen datenverändernder Anonymisierungsverfahren auf die Analyse von Mikrodaten. IAW-Forschungsberichte, Nr. 66, hrsg. v. Institut für Angewandte Wirtschaftsforschung Tübingen. Tübingen 2006.

G 1

## Erhebungsarten in der amtlichen Statistik



Die amtliche Statistik ist bei ihren Erhebungen in besonderem Maße auf das Vertrauen und die Mitarbeit der befragten Bürgerinnen und Bürger sowie der befragten Institutionen angewiesen. Deshalb gehört es zu den Zielen der statistischen Ämter, durch möglichst geringe Belastung der Auskunftgebenden, eine breite Akzeptanz zu gewährleisten. Um dies zu erreichen, wird für die jeweilige Statistik stets die geeignetste – und nach Möglichkeit auch belastungsärmste – Erhebungsmethode gewählt. Andererseits müssen die in der amtlichen Statistik eingesetzten Methoden eine möglichst weitgehende Repräsentativität der Ergebnisse sicherstellen. Aus diesem Grund werden grundsätzlich nur solche Erhebungs- und Hochrechnungsverfahren verwendet, die wirklichkeitsnahe Rückschlüsse auf die Gegebenheiten in der darzustellenden Grundgesamtheit gewährleisten.

Bei Vollerhebungen gehen Informationen von allen Einheiten der Grundgesamtheit ein

Bei Vollerhebungen werden die benötigten Informationen von allen Einheiten der Grundgesamtheit (z. B. allen Unternehmen des produzierenden Gewerbes) eingeholt und ausgewertet. Vollerhebungen bieten in der Regel aussagekräftige Ergebnisse in tiefer fachlicher und regionaler Gliederung.

Häufig zwingen jedoch neben dem Akzeptanzkriterium auch Ansprüche im Hinblick auf die Aktualität sowie Kostenaspekte dazu, eine Statistik nicht als Vollerhebung, sondern in Form einer Teilerhebung durchzuführen. Unabhängig hiervon bleibt in methodischer Hinsicht die Vollerhebung notwendige Voraussetzung für die Schaffung einer vollständigen Grundlage für die Auswahl von Erhebungseinheiten bei Teilerhebungen. Für diese muss sichergestellt sein, dass sie weitgehend alle erhebungsrelevanten Einheiten umfasst und dass die bei der Auswahl zu berücksichtigenden Charakteristika der Auswahlseinheiten nachgewiesen werden. Nur wenn diese Voraussetzungen erfüllt sind, ist eine repräsentative Auswahl der Erhebungseinheiten möglich.

Die Rahmenbedingungen erlauben nur in Einzelfällen die Durchführung von Vollerhebungen

Werden die Erhebungseinheiten aus der Grundgesamtheit nach dem Zufallsprinzip ausgewählt, so wird von einer Zufallsstichprobe gesprochen. Ein Beispiel ist der Mikrozensus. Auf diese Erhebung wird im dritten Teil des Beitrags in der Aprilausgabe des Statistischen Monatshefts detaillierter eingegangen werden.

Bei Stichproben erfolgt die Auswahl der Erhebungseinheiten zufällig

Obwohl bei Stichprobenerhebungen nur ein Teil der Grundgesamtheit einbezogen wird, ermöglichen sie durch die zufällige Auswahl der Erhebungseinheiten statistisch gesicherte Rückschlüsse auf die Gesamtheit. Die Zuverlässigkeit dieser auf die Grundgesamtheit hochgerechneten Ergebnisse kann mithilfe der Wahrscheinlichkeitsrechnung abgeschätzt werden.

Die Genauigkeit der Ergebnisse bei Stichprobenerhebungen ist im Wesentlichen abhängig von

- dem Umfang und der Struktur der Auswahlgrundlage (Grundgesamtheit),
- dem Stichprobenumfang,
- dem Auswahlverfahren,
- der Schätzmethodik und
- dem Hochrechnungsverfahren.

Das Problem von Stichprobenfehlern steigt mit dem Grad der Regionalisierung

Diese Kriterien haben zugleich Einfluss auf das Regionalisierungspotenzial der aus den Erhebungen gewonnenen Daten. Generell kann festgestellt werden:

- Stichprobenfehler fallen umso geringer aus, je homogener die Grundgesamtheit und je größer der Stichprobenumfang sind und
- bei gegebenem Stichprobenumfang nehmen mit zunehmender Regionalisierung die Problematik von Stichprobenfehlern zu und die Repräsentativität der erhobenen Daten ab.

Bei fehlender Auswahlgrundlage erfolgt willkürliche Auswahl der Erhebungseinheiten

Ist die Durchführung einer Zufallsstichprobe aus methodischen oder anderen Gründen nicht realisierbar, werden Auswahlverfahren herangezogen, bei denen bewusst eine Teilmasse der Grundgesamtheit von der Erhebung ausgeschlossen wird.

Häufig werden solche Teilmassen nach einem Ausschlusskriterium festgelegt und

von der Grundgesamtheit „abgeschnitten“. Solche Abschneideverfahren beruhen auf dem Konzentrationsprinzip; d. h., in der Regel werden hier nur die „großen“ Einheiten der Gesamtheit in die Erhebung einbezogen, weil sie den größten Beitrag zu den relevanten Merkmalen liefern. Teilerhebungen mit Abschneidegrenze sind insbesondere dann ein probates Auswahlverfahren, wenn die Ergebnisse einer Erhebung maßgeblich durch relativ wenige Einheiten mit hohen Merkmalswerten bestimmt sind. So werden etwa bei Erhebungen des produzierenden Gewerbes oder der Einzelhandelsstatistik lediglich die „größeren“ Erhebungseinheiten oberhalb einer festgelegten „Abschneidegrenze“ befragt. Auf solche Erhebungen wird insbesondere bei den im Rahmen des dritten Teils dieses Beitrages dargestellten Wirtschaftsstatistiken näher eingegangen.

Dazu gehören das Abschneideverfahren, ...

Das Abschneideverfahren ist sowohl gegenüber Zufallsstichproben als auch gegenüber Vollerhebungen nicht nur hinsichtlich der Belastung der Befragten vorteilhaft. So genügt bei hinreichender Konzentration auf große Einheiten auch im Vergleich zur Zufallsstichprobe ein vergleichsweise geringerer Auswahlsatz, was nicht nur die Kosten reduziert, sondern auch eine schnellere

### Stichprobenfehler

Die Ergebnisse von Stichprobenerhebungen stimmen i. d. R. nicht genau mit den Ergebnissen einer entsprechenden (idealen) Vollerhebung überein. Vielmehr weichen die Ergebnisse zufallsbedingt davon mehr oder weniger stark ab. Bei der Verwendung von Stichprobenergebnissen muss in der statistischen Praxis daher stets ihr Fehlerbereich berücksichtigt werden. Bei Stichproben, die nach dem Zufallsprinzip gezogen werden, lässt sich anhand mathematischer Methoden die Breite des Fehlerbereichs, dem „Vertrauens-“ oder „Konfidenzintervall“, ermitteln. Der „wahre Wert“ liegt mit einer bestimmten Wahrscheinlichkeit innerhalb dieses Intervalls.

Bereitstellung der Ergebnisse ermöglicht. Ein Nachteil von Teilerhebungen nach dem Abschneideverfahren ist, dass sie keine Aussagen über die „kleineren“ Einheiten unterhalb der Abschneidegrenze ermöglichen. Daher sind sie im streng methodischen Sinne auch nur bedingt repräsentativ für die Gesamtheit. Teilerhebungen mit Abschneidegrenze werden in der amtlichen Statistik daher vorwiegend zur Beobachtung kurzfristiger zeitlicher Entwicklungen in den Wirtschaftsstatistiken eingesetzt (Konjunkturstatistiken), während die Analyse von längerfristigen (strukturellen) Veränderungen häufig Vollerhebungen erfordert (Strukturstatistiken).

... die typische Auswahl ...

Ein weiteres Verfahren zur Ausgestaltung von Teilerhebungen in der amtlichen Statistik ist die „typische Auswahl“. Bei dieser Art der Auswahl werden nach festgelegten Kriterien solche Fälle für die Erhebung ausgewählt, bei denen davon ausgegangen werden kann, dass sie ein aussagekräftiges Abbild der zu beschreibenden Wirklichkeit abgeben. Die typische Auswahl wird angewendet, wenn die Vielfalt und Komplexität der zu beschreibenden Merkmale es unmöglich machen, eine Auswahlgrundlage für eine Zufallsauswahl zu erstellen. Dies gilt beispielsweise für den Bereich der Preisstatistik.

... und die Quotenauswahl

Neben dem genannten Verfahren findet in der amtlichen Statistik auch die Quotenauswahl oder Quotenstichprobe Anwendung. Bei der Quotenauswahl wird die Zusammensetzung der Stichprobe so gesteuert, dass für bestimmte Merkmalsausprägungen Anteilswerte der Grundgesamtheit (Quoten) vorgeschrieben werden. Dieses Auswahlverfahren enthält damit – wie das Abschneideverfahren – willkürliche (sub-

jektive) Momente. Es wird in der amtlichen Statistik kaum genutzt. Wenn es eingesetzt wird, dann nur bei Erhebungen, bei denen sich eine zufallsgesteuerte Auswahl nicht mit vertretbarem Aufwand in die Praxis umsetzen lässt. Ein Beispiel ist die Einkommens- und Verbrauchsstichprobe.

Bei Teilerhebungen, die unter Anwendung des Abschneideverfahrens durchgeführt werden, wird das Regionalisierungspotenzial wesentlich von der Höhe der Abschneidegrenze beeinflusst. Je höher die Abschneidegrenze gesetzt wird, umso weniger Einheiten gelangen in die Auswahl. Gleichzeitig nimmt die Geheimhaltungsproblematik zu, sobald die Erhebungsergebnisse regionalisiert und/oder fachlich gegliedert werden. Daneben sinkt die Adäquanz der Daten, weil die willkürliche Auswahl der Einheiten zwar auf höheren Aggregationsniveaus zwar einigermaßen „repräsentativ“ ist, nicht aber für die regionalisierten Teilgesamtheiten. Diese Problematik stellt sich analog für die Erhebungen, bei denen die Quotenauswahl oder die typische Auswahl Verwendung findet.

Besteht bei einer Teilerhebung der Wunsch, die Daten nicht nur regional, sondern auch fachlich gegliedert auszuweisen, entsteht zwischen diesen beiden Gliederungsdimensionen eine Konkurrenzsituation. Bei zunehmender fachlicher Gliederungstiefe nimmt das Regionalisierungspotenzial der Erhebungsergebnisse, insbesondere wegen der Geheimhaltungsproblematik, ab. Bei Stichprobenerhebungen nimmt darüber hinaus der Stichprobenfehler zu. Daneben sinkt die Repräsentativität der Daten. D. h. letztlich, dass regional und fachlich tief gegliederte Ergebnisse von hoher Qualität nur über Vollerhebungen gewonnen werden können.

Abschneidegrenzen bestimmen das Regionalisierungspotenzial

Bei Teilerhebungen besteht ein „Trade-off“ zwischen fachlicher Gliederungstiefe und dem Regionalisierungspotenzial

### Ausblick

In der April-Ausgabe des Statistischen Monatsheftes werden im dritten Teil dieses Beitrags wichtige Bereiche der amtlichen Statistik aufgezeigt, in denen regionalisierte Erhebungsergebnisse angeboten werden.

Dr. Stefan Weil ist Referent im Referat Analysen, Volkswirtschaftliche Gesamtrechnungen, Veröffentlichungen, Forschungsdatenzentrum.