



04 | 2009

# STATISTISCHE MONATSHEFTE – AUSZUG



Regionaldaten der amtlichen Statistik –

Ein Angebot im Spannungsfeld rechtlicher Restriktionen,  
statistisch-methodischer Beschränkungen und dem  
Bedarf an regional tief gegliederten Daten



#### Fotonachweis

Titelfoto: Jürgen Hammerl

Autorenfoto: Manfred Riege, Fotostudio Jörg, Nassau.

Überarbeitete Auszüge aus den Statistischen Monatsheften 2/2009 bis 4/2009

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



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. Im ersten Teil des Beitrags wird das Angebot des Statistischen Landesamtes an regionalisierten Ergebnissen dargestellt. Der zweite Teil befasst sich mit der grundsätzlichen Frage, welche Möglichkeiten und Grenzen bestehen. Abgeschlossen wird der Beitrag mit Ausführungen zur Regionalisierbarkeit ausgewählter Statistiken.

## Das Angebot des Statistischen Landesamtes

### Was versteht man unter „regionalisierten Daten“?

Regionalisierte Daten haben einen räumlichen Bezug

Unter „regionalisierte Daten“ (auch: „Regionaldaten“ oder „regionalstatistischen Daten“) sind Daten zu verstehen, die einen räumlich-geografischen Bezug aufweisen. Etwas „moderner“ ausgedrückt, bedeutet dies, dass Regionaldaten „georeferenziert“ sind. Die Referenz drückt dann den Bezug zu einer geografischen Einheit aus. Die Daten werden dadurch verortet. In einem engeren Sinne sind unter regionalisierten Daten solche zu verstehen, deren regionaler Bezug eher kleinräumig ist (also z. B. die Kreis-

oder die Gemeindeebene). In diesem Beitrag wird der Begriff im engeren Sinne verwendet.

Oft werden von den Kunden der amtlichen Statistik regional möglichst tief gegliederte Daten gewünscht. Häufig ist dieses Anliegen dann auch noch mit dem Wunsch nach einer möglichst tiefen fachlichen Gliederung der Ergebnisse verbunden. Viele Erhebungen, die von der amtlichen Statistik durchgeführt werden, lassen jedoch – aus Gründen die noch ausgeführt werden – den Ausweis tief gegliederter Ergebnisse nicht oder nur eingeschränkt zu.

Oft Konkurrenzbeziehung zwischen räumlicher und fachlicher Gliederung der Erhebungsergebnisse

Häufig besteht eine Konkurrenzbeziehung (ein „trade-off“) zwischen beiden Dimensionen, sodass bei tieferer fachlicher Gliederung die Regionalisierbarkeit eingeschränkt wird und umgekehrt. Dies gilt insbesondere in Ländern, die relativ klein sind oder eine sehr kleinräumige Verwaltungsgliederung haben. Eine Regionalisierung auf Gemeindeebene stellt sich deshalb in Nordrhein-Westfalen mit 18 Mill. Einwohnern in 396 Gemeinden anders dar als in Rheinland-Pfalz mit 4 Mill. Einwohnern in 2306 Gemeinden.

**Wer sind unsere Kunden und wozu benötigen sie Regionaldaten?**

Regionaldaten dienen als Planungs- und Entscheidungsgrundlage

Regionalisierte Ergebnisse von Erhebungen dienen der vergleichenden Beobachtung und Interpretation von räumlichen Zuständen und Entwicklungen. Häufig werden die Ergebnisse in Tabellen aufbereitet und in Grafiken oder Karten visualisiert. Die Daten dienen allen Gruppen, die am gesellschaftlichen Leben beteiligt sind, als wichtige Informationen. Sie sind Grundlage von Planungen und Entscheidungen; sie können aber auch bei der Erfolgskontrolle und der Schwachstellenanalyse eingesetzt werden.

Hauptkunden sind Entscheidungsträger auf kommunaler und auf Landesebene sowie in Unternehmen und die Forschung

Das Spektrum potenzieller Kunden reicht von den Planern und Entscheidern auf kommunaler und auf Landesebene, in Unternehmen und Verbänden bis hin zur wirtschafts- und sozialwissenschaftlichen Regionalforschung.

Für die kommunale Wirtschaftsförderung ist die Verfügbarkeit von Regionaldaten notwendige Voraussetzung für eine effiziente Arbeit. In den Kommunen wurden diese

Daten in jüngster Vergangenheit aber auch verstärkt für Aufgaben benötigt, die von übergeordneten Ebenen (Land, Bund, Europäische Union) an die Gemeinden hergetragen wurden. So sehen sich die Gemeinden häufig mit wachsenden Anforderungen nationaler und europäischer Städtepolitik konfrontiert. Andererseits werden die Städte verstärkt als „Wachstumsmotoren“ für die bei den EU-Gipfeln in Lissabon (2000) und Göteborg (2001) eingeleiteten Strategien für Wachstum und Beschäftigung („Lissabon-Strategie“) und Umwelt („Göteborg-Strategie“) gesehen. Auch vor dem Hintergrund der strategischen Leitlinien der EU-Kohäsionspolitik erhöhte sich die Notwendigkeit für die Kommunen, ihre Planungen und Entscheidungen auf eine geeignete statistische Basis zu stellen. Der vierte Kohäsionsbericht der EU-Kommission bezieht seine Daten im Wesentlichen aus dem „Urban Audit“ (siehe Textkasten Urban Audit).<sup>1)</sup>

**Urban Audit**

Das Urban Audit ist ein europäischer Städtevergleich, an dem sich 348 große und mittlere Städte der Europäischen Union sowie Norwegens, der Schweiz und der Türkei, darunter 40 deutsche Städte, beteiligen. Die Gesamtfederführung liegt beim Statistischen Amt der Europäischen Gemeinschaft (EUROSTAT), für Deutschland bei der KOSIS-Gemeinschaft Urban Audit, die mit dem Statistischen Bundesamt und den Statistischen Landesämtern, die sich im Netzwerk Stadt- und Regionalstatistik abstimmen, zusammenarbeitet.

Mehr als 300 Merkmale der teilnehmenden Städte, über 150 Merkmale der städtischen Agglomerationen („Larger Urban Zones“) und 30 Merkmale der eigens für das Vorhaben gebildeten Stadtgebietsteile beschreiben die Lebensqualität in diesen städtischen Räumen. Zusammen mit einer fast ebenso großen Zahl von Indikatoren stehen für die am deutschen Urban Audit teilnehmenden Städte und Regionen im Rahmen eines interaktiven Angebots Daten im Internet zum Abruf und zur Auswertung in Form von Tabellen, Grafiken und Karten bereit.<sup>1)</sup>

<sup>1)</sup> Weitere Informationen finden Sie unter <http://www.statistik.nuernberg.de/urban-audit> und [www.urbanaudit.org](http://www.urbanaudit.org).

<sup>1)</sup> Vgl. Trutzel, K.: Entwicklungen in der amtlichen Statistik: Anforderungen städtebezogener Statistik, Vortrag vor dem Ausschuss Regionalstatistik der Deutschen Statistischen Gesellschaft in Düsseldorf (28. und 29. Januar 2008).

**T 1** Nutzer von Regionaldaten

Kunde	Aufgabe (Beispiel)	Hierfür genutzte statistische Quellen
Landesplanung, Planungsgemeinschaften	Landesentwicklungsprogramm, Raumordnungsbericht	Bevölkerungs-, Sozial-, Wirtschafts-, Landwirtschafts-, Verkehrs-, Umweltstatistiken
Verbandsgemeinden, Gemeinden	Bauleitplanung und Flächennutzungsplanung	Bevölkerungsstatistiken, -projektionen, Flächenerhebung
Kreisfreie Städte, Landkreise	Kindertagesstättenplanung	Bevölkerungsstatistiken, Statistik der Kindertagesbetreuung, Projektionen
Kreisfreie Städte, Landkreise, Verbandsgemeinden	Schulentwicklungsplanung	Bevölkerungs- und Schulstatistiken, Projektionen
Kreisfreie Städte, Landkreise	Pflegestrukturplanung	Bevölkerungsstatistiken, Statistik der Pflegeeinrichtungen, Projektionen
Forschungsinstitute	Schuldenmonitoring	Finanzstatistiken
Wirtschaftsverbände, Kammern, Unternehmen	Regionalstrukturanalysen, Konjunkturbeobachtung	Statistiken der Inlands- und Auslandsumsätze, Unternehmensgründungen, Insolvenzen
Verkehrsbetriebe	Betriebsmittelplanung	Schulstatistik, Behindertenstatistik

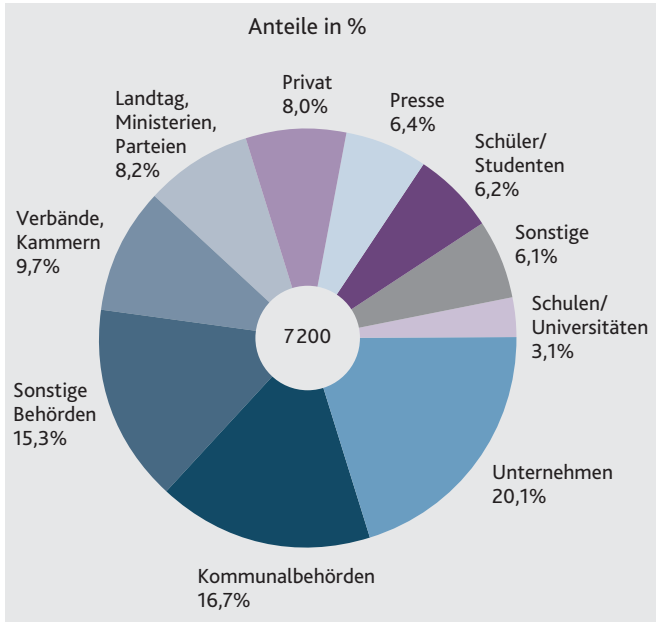
In Unternehmen sind Regionaldaten oft Grundlage von wichtigen Entscheidungen. So ist z. B. in Handelsunternehmen zur Abschätzung und Projektion des Absatzes in verschiedenen Marktregionen die Kenntnis über Niveau und Entwicklung der regionalen Kaufkraft von großer Bedeutung. Kenntnisse über das regionale Angebot an qualifizierten Arbeitskräften sind Grunddaten unternehmerischer Standortentscheidungen. In der Wirtschafts- und in der Sozialforschung (z. B. Marktforschung) bilden Regionaldaten die Grundlage für empirische Analysen. Tabelle 1 zeigt Bei-

spiele für die Nutzungsbereiche regionalstatistischer Daten.

Die bedeutsamsten Nachfragergruppen sind demnach Unternehmen, die Kommunalverwaltungen sowie sonstige Behörden. Die drei Bereiche machten mit Anteilen von 20%, 17% und 15% in der Summe über die Hälfte der Nachfragen an den Auskunftsdienst des Statistischen Landesamts im Jahr 2008 aus (siehe Grafik 1). Aber auch die Anfragen aus dem Landtag, den Ministerien und den Parteien sind mit einem Anteil von über 8% quantitativ bedeutsam.

Über die Hälfte der Anfragen kommt von Unternehmen, Kommunalverwaltungen und sonstigen Behörden

**G 1** Anfragen an den Auskunftsdienst 2008 nach Kundengruppe



Kundenorientierung erfordert kundenspezifische Auswertung und Aufbereitung der Daten

Ein kundenorientiertes Angebot erfordert über die Bereitstellung von Rohdaten hinaus, dass die Ergebnisse der Erhebungen verstärkt auch kundenspezifisch ausgewertet, kommentiert sowie in Form von Tabellen, Grafiken und Karten aufbereitet und visualisiert werden.

**Vielfältiges Datenangebot**

Die regionalstatistische Angebotspalette des Statistischen Landesamtes ist vielseitig. Sie umfasst traditionelle Print-Produkte wie Statistische Berichte, welche die Ergebnisse in standardisierter Form (zumeist in Tabellen) darbieten, ebenso wie interaktive Angebote auf optischen Speichermedien und im Internet.

Zum regionalstatistischen Angebot gehören Statistische Berichte, ...

Eine wichtige Print-Veröffentlichung des Landesamtes mit regionalstatistischen Daten ist die Reihe „Kreisfreie Städte und Landkreise – Ein Vergleich in Zahlen“. Sie bietet auf der Kreisebene mannigfaltige Informationen zu wirtschaftlichen, soziokulturellen und gesellschaftlichen Gegebenheiten, und zwar in Form von Tabellen, Karten und Grafiken. Diese Publikation wird jährlich aktualisiert.

... Statistische Analysen, Kreisübersichten und Kartensammlungen

Dem Wunsch der Kunden nach analytischen Aufbereitungen und grafischer Umsetzung der Ergebnisse entsprechend, wurde in den vergangenen Jahren das Angebot an umfassenderen statistischen Auswertungen

**G 2** Regionalisierte Ergebnisse statistischer Erhebungen in statistischen Analysen



G 3 Internet-Angebot „Meine Heimat“



ausgeweitet. Auswertungen dieser Art erscheinen seit 2006 zumeist in der Reihe „Statistische Analysen“. Dort werden Ergebnisse nicht nur kommentiert, sondern auch Ursachenanalysen vorgenommen.

Ergänzt wird das Angebot an Produkten mit regionalstatistischen Ergebnissen durch Einzelveröffentlichungen. So wurde im Oktober 2008 erstmals ein „Wirtschaftsatlas“ mit einer umfangreichen Sammlung kommentierter regionalstatistischer Karten aufgelegt.<sup>2)</sup>

Die meisten Produkte, die in gedruckter Form angeboten werden, können auch im Internet eingesehen oder auf den eigenen Rechner heruntergeladen werden. Im Internetangebot des Statistischen Landesamtes findet sich seit einigen Jahren auch ein breites Spektrum an regionalstatistischen Angeboten, die z.T. interaktiv gestaltet sind. Zu nennen sind hier insbesondere die Rubriken „Meine Heimat“ und die „Online-Datenbank“.

Print-Produkte auch im Internet verfügbar

„Meine Heimat“ richtet sich an die breite Öffentlichkeit. Mittels übersichtlicher Tabellen und Grafiken können sich die Bürgerinnen und Bürger über die wirtschaftlichen, sozialen und politischen Gegebenheiten ihrer Gemeinde, Verbandsgemeinde oder Stadt, ihres Kreises oder ihres Landes informieren.

Meine Heimat

Die Online-Datenbank enthält statistische Daten über Rheinland-Pfalz und seine Regionen in frei recherchierbarer Form. Sie ist das verkleinerte Abbild des vom Statistischen Landesamt entwickelten, aber nur einem geschlossenen Benutzerkreis zugänglichen, Landesinformationssystems (LIS).

Online-Datenbank

Im Rahmen des Analyseprojekts „Rheinland-Pfalz regional“ wird ein „Datenkompass“ angeboten.

Datenkompass

Dieser Datenkompass enthält zu jeder kreisfreien Stadt und zu jedem Landkreis in thematisch aufgeteilten Modulen regio-

2) Die einzelnen Karten stehen im Internet zum kostenfreien Download bereit unter <http://www.statistik.rlp.de/verlag/wiatlas/download.html>

**G 4** Internet-Angebot „Online-Datenbank“

**G 5** Internet-Angebote „Datenkompass“ und „Datenblätter“

nalstatistische Tabellen zu den Bereichen Bevölkerung, Wirtschaft, Arbeitsmarkt, Tourismus, Innovationsfähigkeit, Gesellschaft und Staat sowie Gebiet, die als PDF-Dokumente ebenfalls über das Internetangebot des Statistischen Landesamtes einer breiten Öffentlichkeit zugänglich

gemacht werden. Zu ausgewählten Merkmalen bzw. Indikatoren werden Karten angeboten. Wesentliches Merkmal der veröffentlichten Tabellen ist, dass die ausgewiesenen Kreiswerte um Rangziffern ergänzt sind, die ein Benchmarking der rheinland-pfälzischen Regionen ermöglichen.



G 6

Internet-Angebot „Gesundheitsberichterstattung“

The screenshot shows the website interface for the 'Gesundheitsberichterstattung Rheinland-Pfalz - Statistische Basisdaten'. The header includes the logo of the 'Statistisches Landesamt RHEINLAND-PFALZ' and navigation links for 'Wahlen', 'Datenerhebung', and 'Forschungsdatenzentrum'. A sidebar on the left lists menu items such as 'Adressenverzeichnisse', 'Meine Heimat', 'Online-Datenbank', and 'Gesundheitsberichterstattung'. The main content area features a photograph of a doctor examining a patient, followed by introductory text about the data provided and a copyright notice for the 'Statistisches Landesamt Rheinland-Pfalz'.

Gesundheitsberichterstattung

Schließlich stellt das Statistische Landesamt im Auftrag des Ministeriums für Arbeit, Soziales, Gesundheit, Familie und Frauen im Internet ausgewählte statistische Basisdaten der Gesundheitsberichterstattung des Landes zur Verfügung, die z. T. auch als Regionalindikatoren genutzt werden können.

**Das Datenangebot des Forschungsdatenzentrums der Statistischen Landesämter**

Forschungsdatenzentrum bietet neue Zugangswege zu Regionaldaten

Für wissenschaftliche Nutzer haben sich seit dem Jahr 2002 mit dem Forschungsdatenzentrum (FDZ) der Statistischen Landesämter weitere Zugangswege zu Regionaldaten eröffnet.<sup>3)</sup> Aufgabe des FDZ ist es, ein breites Spektrum an Einzeldaten der amtlichen Statistik sowie die zugehörigen Metadaten für die wissenschaftliche Nutzung bereitzustellen.

3) Das Statistische Bundesamt hatte bereits im Herbst 2001 ein Forschungsdatenzentrum eingerichtet.

Das FDZ ist als Netzwerk angelegt, dessen Dienstleistungsangebot über 16 regionale Standorte genutzt werden kann. Das Datenangebot umfasst gegenwärtig 69 amtliche Statistiken aus den Bereichen

- Bevölkerungs- und Sozialstatistiken,
- Wirtschaftsstatistiken,
- Finanz- und Steuerstatistiken,
- Rechtspflegestatistiken sowie
- Agrar- und Umweltstatistiken.

In 50 von diesen 69 Statistiken sind Erhebungsergebnisse bis zur Kreisebene verfügbar, in 43 reicht die regionale Gliederungstiefe sogar bis zur Gemeindeebene. Tabelle 2 bietet einen Überblick.

In den meisten verfügbaren Statistiken stehen Regionalergebnisse bereit

Die Nutzung der Daten kann über drei verschiedene Zugangswege erfolgen:

- Standardisierte Mikrodatensätze (Public Use Files bzw. Scientific Use Files),
- Arbeitsplätze für Gastwissenschaftler sowie
- kontrollierte Datenfernverarbeitung.

**T 2** Regionaldaten im Angebot des Forschungsdatenzentrums der Statistischen Landesämter

Fachgebiet	Erhebungen insgesamt	Darunter Regionalisierungstiefe bis ...				
		Land	Regierungsbezirk	kreisfreie Stadt/Landkreis	Gemeinde	Sonstige Gliederungen
Sozialstatistiken	27	22	13	15	8	8
darunter Bevölkerungsstatistiken	8	8	6	7	2	-
Wirtschaftsstatistiken	21	20	19	19	19	-
Finanz-/Steuerstatistiken	8	8	1	7	7	-
Rechtspflegestatistiken	3	3	-	-	-	2
Wahlstatistik	1	1	1	1	1	1
Agrar-/Umweltstatistiken	9	9	8	8	8	-
Insgesamt	69	63	42	50	43	11

Die Verfügbarkeit von Regionaldaten ist bei den verschiedenen Nutzungswegen unterschiedlich

Die Nutzungswege unterscheiden sich sowohl hinsichtlich des Anonymisierungsgrads der Daten als auch in der Art der Bereitstellung (siehe Textkasten Anonymisierungsgrad von Daten).

male enthalten. Fachlich tief gegliederte Merkmale werden in der Regel aggregiert. Tiefere räumliche Abgrenzungen können auf der Basis von Public Use Files deshalb meist nicht vorgenommen werden.

In Public Use Files sind Regionaldaten nur eingeschränkt verfügbar

Die standardisierten Mikrodatensätze können außerhalb der statistischen Ämter genutzt werden. Einer breiten Öffentlichkeit stehen absolut anonymisierte Mikrodaten („Public Use Files“) zur Verfügung. Zum Zwecke der Anonymisierung sind in diesen Datensätzen nur ausgewählte Merk-

Einem wissenschaftlichen Nutzerkreis werden in Form standardisierter Scientific Use Files faktisch anonymisierte Mikrodaten aus diversen amtlichen Statistiken angeboten. Diese Daten verfügen über ein deutlich höheres Informationspotenzial und damit auch Regionalisierungspotenzial als Public Use Files. Sie eignen sich insoweit für ein weites Spektrum wissenschaftlicher Datenanalyse.

Regionalisierungspotenzial bei Scientific Use Files deutlich höher als bei Public Use Files

**Anonymisierungsgrad von Daten**

**Absolut** anonymisiert sind Daten dann, wenn sowohl merkmalsträger- als auch merkmals- und ausprägungsbezogene Informationen so weit reduziert oder verändert sind, dass eine Reidentifikation einzelner Merkmalsträger absolut ausgeschlossen ist.

**Faktisch** anonymisiert sind Daten dann, wenn sowohl merkmalsträger- als auch merkmals- und ausprägungsbezogene Informationen so weit reduziert oder verändert sind, dass eine Reidentifikation einzelner Merkmalsträger nur mit unverhältnismäßig hohem Aufwand möglich ist.

**Formal** anonymisiert sind Daten dann, wenn lediglich merkmalsträgerbezogene Informationen unterdrückt sind.

Neben diesen standardisierten Datenangeboten bietet das FDZ der statistischen Landesämter mit den Gastwissenschaftlerarbeitsplätzen und der kontrollierten Datenfernverarbeitung speziell auf den jeweiligen Datenbedarf zugeschnittene Nutzungsmöglichkeiten amtlicher Statistiken an. Hier können weniger stark anonymisierte Mikrodaten analysiert werden, die dafür eigens in abgeschotteten Bereichen in den statistischen Ämtern bereitgestellt werden.

Höchstes Regionalisierungspotenzial bei Datennutzung an Gastwissenschaftlerarbeitsplätzen und bei kontrollierter Datenfernverarbeitung

An allen Regionalstandorten stehen PC-Arbeitsplätze bereit, an denen faktisch anonymisierte Einzeldaten im geschützten Bereich der amtlichen Statistik von Gastwissenschaftlern ausgewertet werden können. Die faktische Anonymität wird hierbei nicht allein durch die Anonymisierung der Daten (wie bei Scientific Use Files) erreicht, sondern durch spezielle Regelungen des Datenzugangs gewährleistet. Diese Daten dürfen daher wesentlich detailliertere Informationen enthalten (mit entsprechend höherem

Regionalisierungspotenzial) als die bereitgestellten Scientific Use Files.

Die kontrollierte Datenfernverarbeitung erlaubt als einziger Zugangsweg die Analyse formal anonymisierter Originaldaten. Für den Datennutzer besteht hier kein direkter Zugang zu den Daten. Vielmehr stellen sie Auswertungsprogramme bereit, mit denen die statistischen Ämter die Originaldaten auswerten. Nach einer Geheimhaltungsprüfung erhalten die Datennutzer die Ergebnisse der Auswertung.

## Möglichkeiten und Grenzen der Regionalisierung

Räumliche Bezugnahme auf Einheiten der Verwaltungsgliederung

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

T 3

### 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.

Tabelle 3 zeigt die NUTS-Gliederung für das Land Rheinland-Pfalz.

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

Auch andere Strukturräume lassen sich abbilden

- 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“.

Nachweis regionalisierter Erhebungsergebnisse unterhalb der Gemeindeebene bislang noch nicht möglich

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>4)</sup>

### 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.

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

Auskunftspflicht und Datenschutz stellen eine hohe Qualität der Daten sicher

4) 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.

## 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.

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: Einzelangaben über persönliche und sachliche Verhältnisse, die für eine amtliche Statistik gemacht werden, sind gemäß § 16 BStatG geheim zu hal-

ten, 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>5)</sup>

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 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 Schaf-

Bei Vollerhebungen gehen Informationen von allen Einheiten der Grundgesamtheit ein

Die Rahmenbedingungen erlauben nur in Einzelfällen die Durchführung von Vollerhebungen

5) 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 Tabellen-geheimhaltung (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 (Datenveränderung) eine Reidentifikation der Merkmalsträger weitgehend zu verhindern. Informationsre-

duktion 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.

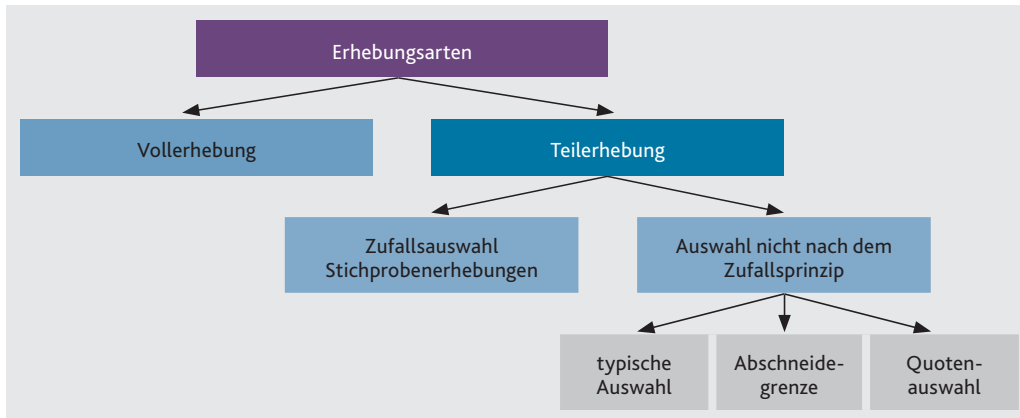
Werden lediglich merkmalsträgerbezogene Informationen wie Name, Anschrift oder z. B. Sozialversicherungsnummer unterdrückt, liegt eine „formale Anonymisierung“ vor. Formal anonymisierte Mikrodaten werden von den FDZ lediglich im Rahmen der kontrollierten Datenfernverarbeitung bereitgestellt. „Faktische 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. Statistischem 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 7 Erhebungsarten in der amtlichen Statistik



fung 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 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.

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.

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.

Das Problem von Stichprobenfehlern steigt mit dem Grad der Regionalisierung

Bei fehlender Auswahlgrundlage erfolgt willkürliche Auswahl der Erhebungseinheiten

Bei Stichproben erfolgt die Auswahl der Erhebungseinheiten zufällig

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 weiter unten detaillierter eingegangen.

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.

**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.

**Abschneideverfahren**

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 am Ende des Beitrages dargestellten Wirtschaftsstatistiken näher eingegangen.

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 Auswahlatz, was nicht nur die Kosten reduziert, sondern auch eine schnellere

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).

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.

Typische Auswahl

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) vorge-schrieben werden. Dieses Auswahlverfahren enthält damit – wie das Abschneideverfahren – willkürliche (subjektive) Momente. Es wird in der amtlichen Statistik kaum genutzt.

Quotenauswahl



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.

Abschneidegrenzen bestimmen das Regionalisierungspotenzial

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.

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

## Regionalisierungspotenzial in den Wirtschaftsstatistiken, im Mikrozensus und im Zensus 2011

### Datenangebot und Regionalisierungspotenzial in den Wirtschaftsstatistiken

Informationsquellen über die Unternehmens- und Betriebsstrukturen

Die Wirtschaftsstatistiken informieren über die Unternehmens- und Betriebsstrukturen in den einzelnen Wirtschaftsbereichen („Strukturstatistiken“) sowie über die konjunkturelle Entwicklung („Konjunkturstatistiken“).

Zu den Strukturstatistiken gehören z. B.

- die Strukturserhebungen in der Landwirtschaft,
- die Monats- und Jahresberichte im verarbeitenden Gewerbe,
- die Investitionserhebungen im verarbeitenden Gewerbe sowie im Bau- und im Ausbaugewerbe,

- die Ergänzungserhebung im Bauhauptgewerbe,
- die Jahresherhebungen im Handel und im Gastgewerbe sowie
- die Strukturserhebung im Dienstleistungssektor.

Ebenso dazu gehören bereichsübergreifende Erhebungen wie

- die Gewerbeanzeigenstatistik,
- die Insolvenzstatistik und
- das Unternehmensregister.

Die Strukturstatistiken werden ergänzt durch monatliche sowie vierteljährliche Erhebungen, die hauptsächlich der laufenden Konjunkturbeobachtung dienen. Zu den Konjunkturstatistiken gehören

Konjunkturstatistiken dienen der Konjunkturbeobachtung

- die Monatsberichte im verarbeitenden Gewerbe und im Bauhauptgewerbe,
- die vierteljährliche Erhebung im Ausbaugewerbe,
- die Monaterhebungen im Handel und im Gastgewerbe sowie
- die sich im Aufbau befindliche vierteljährliche Konjunkturstatistik im Dienstleistungsbereich.

Als Konjunkturindikatoren werden u. a. Angaben über Beschäftigung und Umsätze ermittelt, in Teilbereichen auch Auftragseingänge und -bestände.

Für bestimmte Zwecke werden in größeren Zeitabständen Totalerhebungen durchgeführt

Als Gesamtdarstellungen der wirtschaftlichen Verhältnisse in tiefer regionaler und fachlicher Gliederung sowie als Hochrechnungsrahmen bzw. Auswahlgrundlage für die verschiedenen Teil- und Repräsentativerhebungen sind in größeren Zeitabständen Vollerhebungen vorgesehen (z. B. die Arbeitsstättenzählung). Tabelle 4 verschafft einen Überblick über das Angebot an regionalisierten Daten in den Wirtschaftsstatistiken. Die Ergebnisse werden jeweils in Statistischen Berichten veröffentlicht. Diese stehen kostenfrei im Internet unter <http://www.statistik.rlp.de/verlag/gesamt/index.html> zur Verfügung.

Für einzelne Wirtschaftsbereiche unterschiedliches Regionalisierungspotenzial

Bei den Strukturhebungen in der Landwirtschaft lassen sich lediglich für die allgemein erhobenen Merkmale regionalisierte Ergebnisse ausweisen, nicht jedoch für repräsentativ erhobene Merkmale. Im produzierenden Gewerbe liefern die Statistiken auch regional tief gegliederte Ergebnisse, in den Dienstleistungsbereichen kann keine Regionalisierung erfolgen. In den bereichsübergreifenden Erhebungen lassen sich die Erhebungsergebnisse dagegen durchweg bis auf die Kreisebene regionalisieren.

Bei den Konjunkturstatistiken liegen nur bei Statistiken, die als Teilerhebungen nach dem Abschneideverfahren durchgeführt werden, regionalisierte Ergebnisse vor. Bei Stichprobenerhebungen ist keine Regionalisierung möglich. Dies betrifft die (meisten) Erntestatistiken in der Landwirtschaft wie auch sämtliche Erhebungen im Dienstleistungsbereich.

Regionalisierung in Konjunkturstatistiken

Nach Wirtschaftsbereichen betrachtet ist das Regionalisierungspotenzial bei den Erhebungen im produzierenden Gewerbe durchgängig hoch, in den Agrarstatistiken nur bei Teilerhebungen, die nach dem Abschneideverfahren durchgeführt werden. Bei den Erhebungen in den Dienstleistungsbereichen, bei denen zumeist Stichprobenerhebungen durchgeführt werden, ist das Regionalisierungspotenzial der Erhebungsergebnisse am geringsten. In keiner der aufgeführten Erhebungen ist eine Regionalisierung unterhalb der Landesebene durchführbar.

### Datenangebot und Regionalisierungspotenzial beim Mikrozensus

Der Mikrozensus ist die größte amtliche Haushaltebefragung in Deutschland. Seit 1957 wird im Rahmen dieser Erhebung jährlich 1% der Privathaushalte befragt. Die Stichprobe, deren Auswahl immer noch auf der Volks- bzw. der Gebäude- und Wohnungszählung von 1987 aufsetzt, umfasst für Rheinland-Pfalz rund 18 500 Haushalte. Davon konnten im Jahr 2007 lediglich etwa 15 700 Haushalte mit rund 36 700 Personen befragt werden. Bundesweit waren es etwa 370 000 Haushalte mit rund 820 000 Personen. Das Ausmaß der Ausfälle (falsche Adressen, Haushalte verzogen, Haus abgerissen usw.) – trotz jährlicher Aktualisierung durch die Bautätigkeitsstatistik immerhin rund 15% – zeigt die

Mikrozensus größte amtliche Haushaltebefragung

## T 4

## Ausgewählte Wirtschaftsstatistiken und das Regionalisierungspotenzial ihrer Ergebnisse – Strukturstatistiken

Statistik	Merkmale	Erhebungsumfang <sup>1)</sup> , Erhebungsart und Auswahlverfahren	Maximale Regionalisierungstiefe <sup>2)</sup>
<b>Landwirtschaft</b>			
Strukturerhebungen in der Landwirtschaft	Betriebe, Anbauflächen, Viehbestände, Arbeitskräfte, Sozialökonomik, Pachtpreise	Etwa 25 000 landwirtschaftliche Betriebe Teilerhebungen mit unterschiedlichen Abschneidegrenzen	KS, LK (teilweise G)
<b>Produzierendes Gewerbe</b>			
Investitionserhebung im verarbeitenden Gewerbe	Bruttoanlageinvestitionen (Maschinen, Anlagen, Grundstücke, Gebäude), Mietinvestitionen (Absolutwerte)	2 000 Betriebe Teilerhebung mit Abschneidegrenze (20 Beschäftigte)	KB, KS, LK
Monats- und Jahresbericht im verarbeitenden Gewerbe	Betriebe, Umsatz, Beschäftigte, Entgelte (Absolutwerte)	2 000 Betriebe Teilerhebung mit Abschneidegrenze (20 Beschäftigte)	KB, KS, LK
Ergänzungserhebung im Bauhauptgewerbe	Betriebe, Umsatz, Beschäftigte, Auftragseingang, Arbeitsstunden (Absolutwerte)	3 750 Betriebe Totalerhebung	KB, KS, LK
Investitionserhebungen im Bauhauptgewerbe und im Ausbaugewerbe	Bruttoanlageinvestitionen (Maschinen, Anlagen, Grundstücke, Gebäude), Mietinvestitionen (Absolutwerte)	Jeweils rund 300 Unternehmen Teilerhebungen mit Abschneidegrenze (20 Beschäftigte)	KB, KS, LK
<b>Dienstleistungen</b>			
Jahreserhebung im Handel sowie in der Instandhaltung und Reparatur von Kfz und Gebrauchsgütern	Unternehmen, Beschäftigte, Umsatz, Warenbestände, Aufwendungen, Investitionen, Wareneinsatz, Rohertrag (Absolutwerte)	3 300 Unternehmen Geschichtete 8%-Stichprobe	–
Jahreserhebung im Gastgewerbe	Unternehmen, Beschäftigte, Umsatz, Warenbestände, Aufwendungen, Investitionen, Rohertrag (Absolutwerte)	700 Unternehmen Geschichtete 8%-Stichprobe	–
Strukturerhebung im Dienstleistungssektor	Unternehmen, Umsatz, Beschäftigte, Entgelte, Investitionen, Steuern, Subventionen (Absolutwerte)	7 500 Unternehmen Geschichtete 20%-Stichprobe	–
<b>Bereichsübergreifend</b>			
Gewerbeanzeigenstatistik	Wirtschaftszweig, Art des Betriebs, Grad der Selbstständigkeit	212 Gewerbeämter Totalerhebung der Gewerbean- und -abmeldungen 22 Insolvenzgerichte (Amtsgerichte)	KS, LK
Insolvenzstatistik	Insolvenzverfahren, voraussichtliche Forderungen, Beschäftigte	Totalerhebung der eröffneten und mangels Masse abgelehnten Insolvenzverfahren sowie der angenommenen Schuldenbereinigungspläne	KS, LK
Unternehmensregister	Art der Einheit, Wirtschaftszweig, Beschäftigte, Umsatz	Regelmäßige Einpflege administrativer Dateien Datenaustausch mit statistischen Ämtern und Fachbereichen	KS, LK (Umsatz: –)

1) Erhebungsumfang in Rheinland-Pfalz. – 2) Kreisfreie Städte (KS); Landkreise (LK), Gemeinden (G); Kammerbezirke (KB); keine Regionalisierung möglich (–).

T 5

Ausgewählte Wirtschaftsstatistiken und das Regionalisierungspotenzial ihrer Ergebnisse – Konjunkturstatistiken

Statistik	Merkmale	Erhebungsumfang <sup>1)</sup> , Erhebungsart und Auswahlverfahren	Maximale Regionalisierungstiefe <sup>2)</sup>
<b>Landwirtschaft</b>			
Bodennutzungsstatistiken (Gemüseanbau-, Zierpflanzen-, Baumschul- und Baumobsterhebung)	Anbauflächen von Gemüse, Baumschulen und Baumobst, Stückzahlen von Zierpflanzen	Etwa 600 landwirtschaftliche Betriebe  Teilerhebungen mit unterschiedlichen Abschneidegrenzen (bei Anbauflächen)	KS, LK, (teilweise G)
Erntestatistiken	Erntemengen und Erträge für Feldfrüchte, Gemüse, Obst, Weinmost	Stichprobe auf freiwilliger Basis	-
Weinstatistiken	Bestockte Rebflächen, Weinerzeugung, Weinbestände	Auswertung der von der Landwirtschaftskammer Rheinland-Pfalz geführten Weinbaukartei	AG
<b>Produzierendes Gewerbe</b>			
Monatsbericht im verarbeitenden Gewerbe	Betriebe, Umsatz, Beschäftigte, Exportquote, Auftragseingang, Entgelte, Arbeitsstunden (Absolutwerte)	1 030 Betriebe  Teilerhebung mit Abschneidegrenze (50 Beschäftigte)	KB, KS, LK
Monatsbericht im Bauhauptgewerbe	Betriebe, Umsatz, Beschäftigte, Exportquote, Auftragseingang, Entgelte, Arbeitsstunden (Absolutwerte)	350 Betriebe  Teilerhebung mit Abschneidegrenze (20 Beschäftigte)	KB, KS, LK
Vierteljährliche Erhebung im Ausbaugewerbe	Betriebe, Umsatz, Beschäftigte, Exportquote, Auftragseingang, Entgelte, Arbeitsstunden (Absolutwerte)	310 Betriebe  Teilerhebung mit Abschneidegrenze (20 Beschäftigte)	KB, KS, LK
<b>Dienstleistungen</b>			
Monatserhebung im Kfz-/Einzelhandel	Voll- und Teilzeitbeschäftigte, Umsatz (Messzahlen; nominal und real)	2 200 Unternehmen  Geschichtete 9%-Stichprobe mit Abschneidegrenze (250 000 Euro Jahresumsatz)	-
Monatserhebung im Gastgewerbe	Voll- und Teilzeitbeschäftigte, Umsatz (Messzahlen; nominal und real)	700 Unternehmen  Geschichtete 8%-Stichprobe mit Abschneidegrenze (50 000 Euro Jahresumsatz)	-
Monatserhebung im Großhandel	Voll- und Teilzeitbeschäftigte, Umsatz (Messzahlen; nominal und real)	530 Unternehmen  Geschichtete 9%-Stichprobe mit Abschneidegrenze (1 000 000 Euro Jahresumsatz)	-
Vierteljährliche Konjunkturstatistik im Dienstleistungsbereich (im Aufbau)	Beschäftigte, Umsatz (Messzahlen)	100 Unternehmen aus Primärerhebung;  Teilerhebung mit Abschneidegrenze (15 Mill. Euro Jahresumsatz oder 250 Beschäftigte);  übrige Unternehmen:  Auswertung von Daten der Finanz- und der Arbeitsverwaltung	-

1) Erhebungsumfang in Rheinland-Pfalz. – 2) Kreisfreie Städte (KS); Landkreise (LK), Gemeinden (G); Kammerbezirke (KB); Anbaugebiete und -bereiche (AG); keine Regionalisierung möglich (-).

Notwendigkeit einer neuen Vollerhebung, um die Auswahlgrundlage wieder auf den neuesten Stand zu bringen.

Informationen in tiefer fachlicher Gliederung über Bevölkerung

Der Mikrozensus stellt Informationen bereit in tiefer fachlicher Gliederung über die Bevölkerung und deren

- wirtschaftliche und soziale Lage,
- schulische und berufliche Qualifikation,
- Erwerbsbeteiligung,
- Gesundheit,
- Wohnsituation sowie
- Haushalts- und Familienstrukturen.

Mikrozensus füllt Lücke zwischen Volkszählung ...

Der Mikrozensus dient dazu, in regelmäßigen und kurzen Abständen Eck- und Strukturdaten über die genannten Erhebungsinhalte sowie deren Veränderung zu ermitteln und dadurch die Datenlücke zwischen zwei Volkszählungen zu füllen. Dabei macht die Gestaltung als Mehrthemenumfrage eine größere Zahl von sonst notwendigen zusätzlichen Einzelerhebungen entbehrlich. Dies reduziert die Erhebungskosten und trägt zur Entlastung der Befragten bei. Für eine Reihe kleinerer Erhebungen der empirischen Sozial- und Meinungsforschung sowie der amtlichen Statistik dient der Mikrozensus auch als Hochrechnungs-, Adjustierungs- und Kontrollinstrument.

... und dient als Hochrechnungs-, Adjustierungs- und Kontrollinstrument

Die im Rahmen des Mikrozensus erhobenen Daten ermöglichen

- vernetzte Auswertungen zu unterschiedlichen Themenbereichen,
- personen-, familien- und haushaltsbezogene Einzeldatenauswertungen,
- Strukturuntersuchungen im längerfristigen Vergleich (Längsschnittstudien) sowie
- repräsentative regionalisierte Auswertungen bis zur Ebene von acht regionalen Anpassungsschichten (siehe Textkasten).

Wie auch die Daten aus den Wirtschaftsstatistiken werden die Ergebnisse des Mikrozensus als Statistische Berichte kostenfrei im Internet veröffentlicht.

Die Stichprobe wird als einstufige Klumpenstichprobe (Flächenstichprobe) durchgeführt. Auswahlseinheiten sind demnach Klumpen, hier in der Form künstlich abgegrenzter Flächen (Auswahlbezirke), die sich aus Gebäuden oder Gebäudeteilen zusammensetzen. Alle Personen bzw. Haushalte in einem Auswahlbezirk sind als Erhebungseinheiten zu erfassen.

Die Stichprobe ist des Weiteren so angelegt, dass neben einer fachlichen Schichtung, auf die hier nicht näher eingegangen wird, eine regionale Schichtung erfolgt. Letztere soll die regionale Repräsentativität sicherstellen. Als regionale Schichten dienen bundesweit 201 Raumeinheiten von durchschnittlich etwa 350 000 Einwohnern. Großstädte ab 200 000 Einwohnern und andere Regionen ab 250 000 Einwohnern, die in der Regel ein oder mehrere Kreise umfassen, können eigene regionale Schichten bilden. Regionale Schichtuntergruppen mit mindestens 100 000 Einwohnern werden durch eine entsprechende Anordnung der Auswahlbezirke vor der Auswahl berücksichtigt. Die Art der Auswahl (Sortierung, Zonenbildung und Auswahl pro Zone) gewährleistet für diese Regionen einen schichtungsähnlichen Effekt. Die regionalen Schichten werden mindestens so weit zu 123 sogenannten Anpassungsschichten zusammengefasst, dass durchschnittlich 500 000 Einwohner erreicht werden. Für Rheinland-Pfalz ergeben sich dabei acht solcher Anpassungsschichten. Auf der Ebene der Anpassungsschichten erfolgt schließlich die Hochrechnung.

Mikrozensus ist geschichtete Stichprobe

### Anpassungsschichten im Mikrozensus

Den acht Regionen (Anpassungsschichten des Mikrozensus) sind jeweils folgende Verwaltungsbezirke zugeordnet.

#### Mittelrhein-Ost

Landkreise Altenkirchen (Westerwald), Neuwied, Rhein-Lahn-Kreis und Westerwaldkreis

#### Mittelrhein-West

Kreisfreie Stadt Koblenz, Landkreise Ahrweiler, Cochem-Zell, Mayen-Koblenz, und Rhein-Hunsrück-Kreis

#### Nahe

Landkreise Bad Kreuznach und Birkenfeld

#### Rheinhessen

Kreisfreie Städte Mainz und Worms, Landkreise Alzey-Worms und Mainz-Bingen

#### Südpfalz

Kreisfreie Stadt Landau in der Pfalz, Landkreise Germersheim und Südliche Weinstraße

#### Trier

Kreisfreie Stadt Trier, Landkreise Berncastel-Wittlich, Eifelkreis Bitburg-Prüm, Trier-Saarburg und Vulkaneifel

#### Vorderpfalz

Kreisfreie Städte Frankenthal (Pfalz), Ludwigshafen am Rhein, Neustadt an der Weinstraße und Speyer; Landkreise Bad Dürkheim und Rhein-Pfalz-Kreis

#### Westpfalz

Kreisfreie Städte Zweibrücken, Kaiserslautern und Pirmasens; Landkreise Donnersbergkreis, Kaiserslautern, Kusel und Südwestpfalz

Das Stichprobendesign des Mikrozensus ist zwar so gestaltet, dass mögliche Fehler minimiert und kontrolliert werden können. Dennoch sind Stichprobenerhebungen grundsätzlich mit Zufallsfehlern behaftet. Diese sind darauf zurückzuführen, dass nicht alle Einheiten der Grundgesamtheit befragt werden und die Zusammensetzung der Stichprobe vom Zufall abhängt. Zudem treten bei jeder statistischen Messung Fehler auf, die nicht stichprobenbedingt sind. Diese können begrenzt, aber nicht völlig vermieden werden. Um eine hinreichende Genauigkeit der Erhebung zu gewährleisten, ist zum einen der Auswahlsatz relativ hoch (1% der Haushalte), und zum anderen besteht Auskunftspflicht.

Der Umfang eines zufallsbedingten Stichprobenfehlers lässt sich nicht exakt ermitteln, sondern nur abschätzen. Als Schätzwert dient der Standardfehler, der aus den Einzeldaten der Stichprobe berechnet wird. Bei ausreichend großem Stichprobenumfang ist davon auszugehen, dass der jeweilige Wert der Grundgesamtheit mit einer Wahrscheinlichkeit von rund 68% im Bereich des einfachen und mit einer Wahrscheinlichkeit von rund 95% im Bereich des zweifachen Standardfehlers um den hochgerechneten Wert liegt. Nach vorläufigen Berechnungen geht der einfache relative Standardfehler für hochgerechnete Quartalsergebnisse des Mikrozensus von unter 20 000 Personen über 15% hinaus. Bei Jahresergebnissen ist dies bei hochgerechneten Werten, die auf weniger als auf 5 000 Personen zutreffen, der Fall.

Diese Zusammenhänge sind der Grund dafür, dass die Repräsentativität regionalisierter Ergebnisse des Mikrozensus für Rheinland-Pfalz unterhalb der Ebene der acht Anpassungsschichten stark eingeschränkt ist. Dies

Stichproben mit zufälligen und systematischen Fehlern behaftet

Ergebnisse unterhalb der Ebene der Anpassungsschichten nicht hinreichend repräsentativ

**T 6** Mikrozensus – Erhebungseinheiten 2007

Gebiet	Haushalte	Bevölkerung		
		insgesamt	darunter	und zwar
			Erwerbs-tätige	Selbst-ständige
Anzahl				
Rheinland-Pfalz	15 678	36 673	14 719	1 655
Anpassungsschicht Trier	2 025	4 796	2 003	226
Kreisfreie Stadt Trier	529	1 054	432	39

soll an einem Beispiel verdeutlicht werden: Betrachtet wird die Zahl der Selbstständigen in Rheinland-Pfalz, in der Anpassungsschicht Trier und in der Stadt Trier im Jahr 2007.

In jenem Jahr wurden in Rheinland-Pfalz 15 678 Haushalte befragt. In diesen lebten 36 673 Personen, von denen wiederum 1 655 Personen angaben, eine selbstständige Erwerbstätigkeit auszuüben. Von den befragten Haushalten waren 2 025 der Anpassungsschicht Trier zuzuordnen. In diesen Haushalten lebten 4 796 Personen; davon waren 226 selbstständig. Von diesen Haushalten wiederum befanden sich 529 in der Stadt Trier. Dort lebten 1 054 Personen, davon 39 Selbstständige.

Werden diese Werte mit den (für die einzelnen regionalen Ebenen unterschiedlichen) Hochrechnungsfaktoren hochgerechnet, so ergibt sich für Rheinland-Pfalz ein Wert von 205 800 Selbstständigen („Punktschätzer“), für die Anpassungsschicht Trier ein Wert von 27 600 Selbstständigen und für die Stadt Trier ein Wert von 4 400 Selbstständigen. Näherungsweise liegt der relative Standardfehler für die genannten Werte bei 2,6% auf der Landesebene und bei 7% auf der Ebene der Anpassungsschicht. Auf der Kreisebene

**T 7** Mikrozensus – Ergebnisse 2007

Gebiet	Selbstständige				
	Erhebungs-einheiten	Punkt-schätzer	Relativer Standard-fehler <sup>1)</sup>	95%-Vertrauens-bereich <sup>1)</sup>	
				Unter-grenze	Ober-grenze
Anzahl		%	Anzahl		
Rheinland-Pfalz	1 655	205 800	2,6	195 312	216 288
Anpassungsschicht Trier	226	27 600	7,0	23 813	31 387
Kreisfreie Stadt Trier	39	4 400	17,5	2 891	5 909

1) Näherungsweise Ermittlung.

erreicht der relative Standardfehler einen Wert von 17,5%.

Der unbekannt „wahre“ Wert der Zahl der Selbstständigen kann demnach auf Landesebene mit einer Wahrscheinlichkeit von 95% zwischen 195 312 und 216 288, auf Ebene der Anpassungsschicht zwischen 23 813 und 31 387 bzw. auf Kreisebene zwischen 2 891 und 5 909 liegen (Punktschätzer  $\pm$  1,96-facher Standardfehler). Soll der Nachweis (auf Kreisebene) mit noch höherer Wahrscheinlichkeit abgesichert werden, so wird die Aussage noch unpräziser. Bei Vorgabe eines 99-prozentigen Sicherheitsniveaus liegt die Zahl der Selbstständigen in der Stadt Trier zwischen 2 400 und 6 400 (Punktschätzer  $\pm$  2,58-facher Standardfehler).

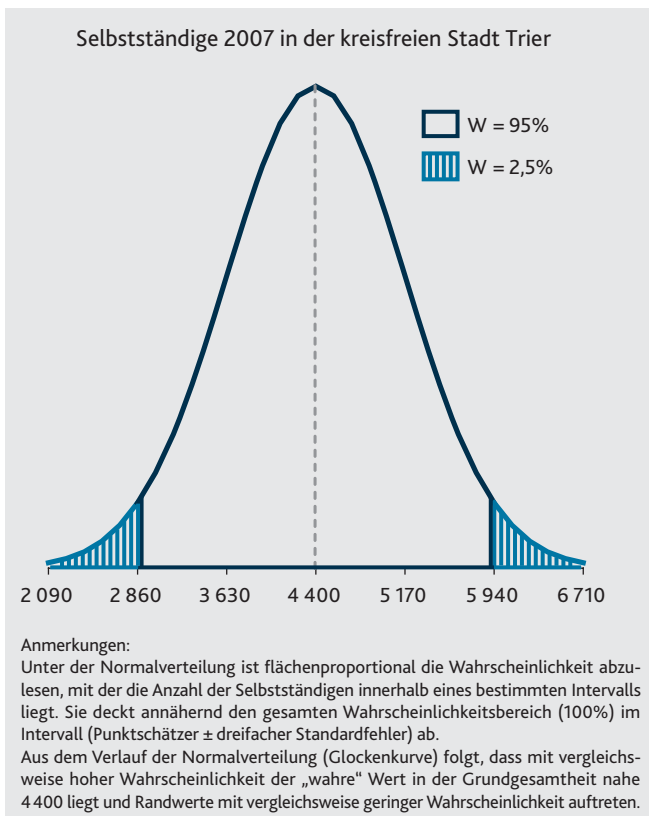
Die Grafik 8 belegt, dass fachlich differenzierten Nachweisen in tiefer regionaler Gliederung durch die Stichprobenfehler enge Grenzen gesetzt werden.

Die Bandbreite, innerhalb derer z. B. die Zahl der Selbstständigen in der Stadt Trier bei Vorgabe eines 95-prozentigen Sicherheitsniveaus schwanken ( $4\,400 \pm 1\,500$  Personen), ist so groß, dass die hochgerechnete Zahl nicht mehr als Planungs- und Entscheidungsgrundlage genutzt werden kann. Selbst

Fachlich differenzierte Nachweise in tiefer regionaler Gliederung nur sehr eingeschränkt möglich

G 8

Mikrozensus – Stichprobenfehler



wenn also auf tiefer regionaler Ebene eine Repräsentativität durch das Stichprobendesign eigentlich gewährleistet wäre, lassen sich die hochgerechneten Werte wegen hoher Standardfehler nicht mehr sinnvoll verwenden.

Probleme bei Regionalisierung

Daher ist neben der mangelnden Repräsentativität der Mikrozensusergebnisse unterhalb der regionalen Ebene der Anpassungsschichten die Nachweisbarkeit von Ergebnissen für kleine Teilmassen zusätzlich durch den hohen Standardfehler eingeschränkt. Der Erhebung vernetzter Strukturdaten im Rahmen des Zensus 2011 kommt deshalb große Bedeutung zu. Auf diesen wird im nächsten Abschnitt eingegangen.

6) Für weitere Informationen zum Zensus 2011 siehe Schneider, D.: Zensus 2011 – Erhebungsablauf und Aufgaben der Kommunen bei der nächsten Volks-, Gebäude- und Wohnungszählung. In Statistische Monatshefte Rheinland-Pfalz 2/2008, S. 129 ff.

Regionalisierungspotenzial beim Zensus 2011

Mit dem für das Jahr 2011 anstehenden Zensus erfolgt ein methodischer Paradigmenwechsel: Es wird keine Volkszählung im herkömmlichen Sinne durchgeführt, also nicht mehr alle Haushalte befragt. Vielmehr wird bei diesem Zensus erstmals eine Kombination aus drei verschiedenen Verfahren angewandt. Dazu gehören

- eine Auszählung von Registern der Meldebehörden, der Bundesagentur für Arbeit sowie der Personalstellen öffentlich Bediensteter,
- primärstatistische Erhebungen (eine Gebäude- und Wohnungszählung, eine Erhebung in Gemeinschafts- und Anstaltsunterkünften und eine Haushaltstichprobe) sowie
- darauf aufsetzende maschinelle Verfahren zur Haushaltgenerierung.<sup>6)</sup>

Kombination aus Registerauszählung, primärstatistischen Erhebungen und maschinellen Verfahren

Neben der Feststellung amtlicher Einwohnerzahlen für Bund, Länder und Gemeinden hat der Zensus das Ziel, vernetzte Strukturdaten in tiefer fachlicher und regionaler Gliederung bereitzustellen und damit eine verlässliche Basis auch für planerische Belange zu schaffen. Das betrifft insbesondere die Fortschreibung der Bevölkerung sowie des Gebäude- und Wohnungsbestandes, aber auch die Bereitstellung einer aktuellen Auswahlgrundlage für Teilerhebungen (wie z. B. den Mikrozensus).

Zensus soll amtliche Einwohnerzahlen als verlässliche Basis für weitere statistische Aufgaben bereitstellen

Vor diesem Hintergrund stellt sich die Frage, ob die Möglichkeit, kleinräumig regionalisierte Erhebungsergebnisse zu erzielen, im Vergleich zum traditionellen Verfahren gleichwertig ist oder aber Einschränkungen in Kauf genommen werden müssen.



## T 8

## Gemeindestruktur 2006 in ausgewählten Ländern

Land	Gemeinden			Bevölkerung		
	insgesamt	darunter über 10 000 Einwohner		insgesamt	darunter in Gemeinden über 10 000 Einwohner	
	Anzahl		%	Anzahl		%
Baden-Württemberg	1 110	250	22,5	10 738 753	7 334 562	68,3
Hessen	426	171	40,1	6 075 359	4 684 761	77,1
Nordrhein-Westfalen	396	343	86,6	18 028 745	17 603 245	97,6
Saarland	52	40	76,9	1 043 167	949 165	91,0
Rheinland-Pfalz						
Gemeindeebene	2 306	45	2,0	4 052 860	1 662 040	41,0
Verbandsgemeindeebene	212	163	76,9	4 052 860	3 633 953	89,7

Aus dem methodischen Ansatz folgt, dass bei Auswertung der Registerquellen und der Gebäude- und Wohnungszählung zwar tief regionalisierte Ergebnisse erzielt werden können. Allerdings stellt sich das Problem einer zieladäquaten Konzeption der ergänzenden Haushaltstichprobe. Über diese soll einerseits die Qualität der aus den Melde-registern gewonnenen Daten gesichert werden. Andererseits müssen aus der Haushal- tebefragung Strukturdaten ermittelt werden, die sich nicht aus Registern ableiten lassen. Letzteres betrifft vorwiegend Informationen zur Erwerbstätigkeit und zum Bildungsstand der Bevölkerung.

Verschiedene  
Vorschläge  
für das Stich-  
probendesign

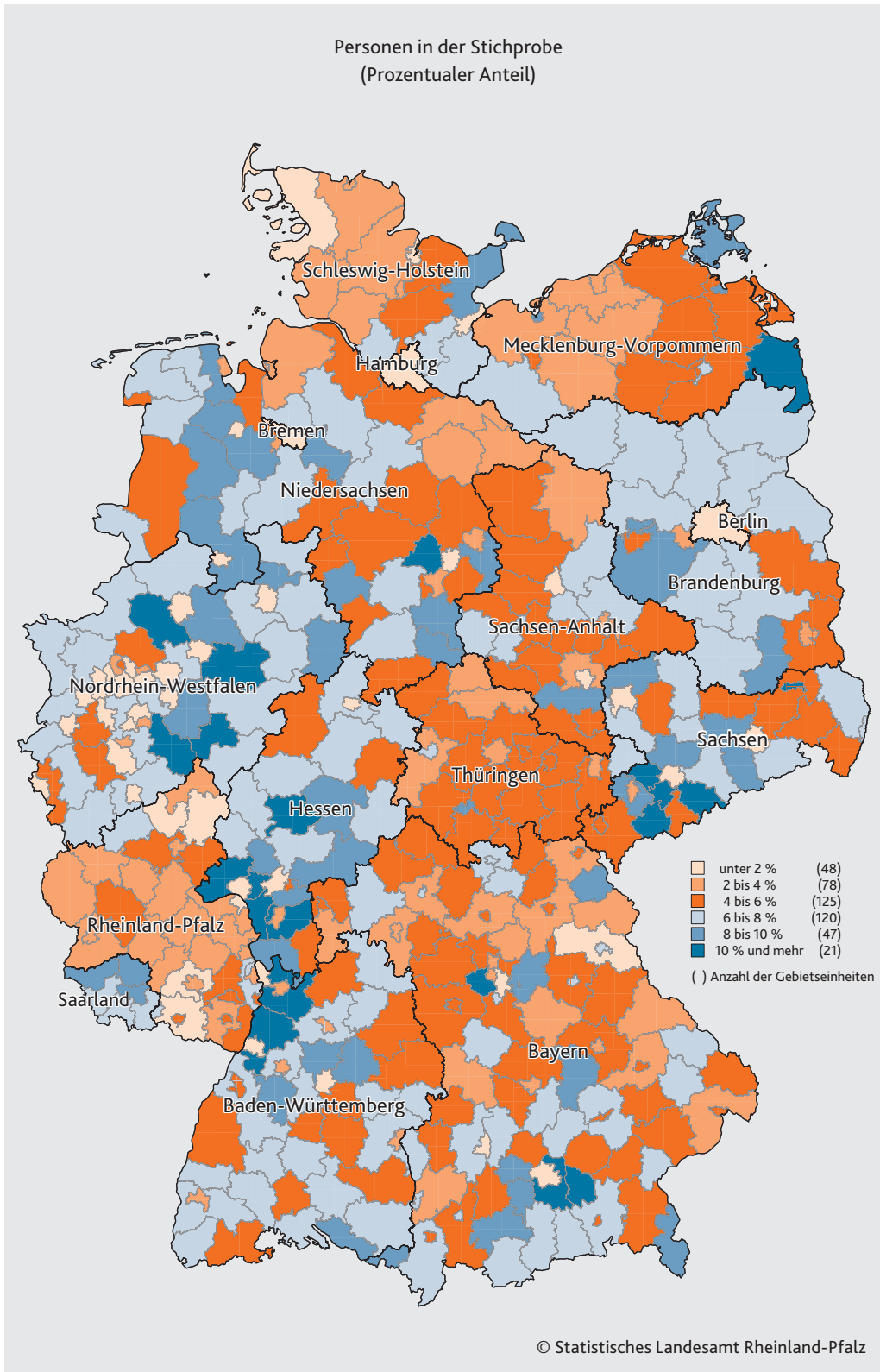
Bezüglich der Gestaltung dieser Stich- probe wurden verschiedene Vorschläge in den Raum gestellt, die sowohl in der amtlichen Statistik als auch auf politischer Ebene intensiv diskutiert werden. Eine Ursprungsvariante sieht eine Befragung hauptsächlich von Gemeinden mit mehr als 10 000 Einwohnern mit gleicher Anzahl von Erhebungseinheiten je Großgemeinde (550 Adressen) und einer disproportionalen Ausgestaltung des Stichproben- designs hinsichtlich der Adressgröße vor, um damit die Qualität der amtlichen Ein-

wohnerzahlen zu verbessern. Der Grund dieser Vorgehensweise ist die Vermutung, dass in Großadressen mehr Fehlbestände und/oder Karteileichen vorzufinden sind. Für die statistischen Zusatzmerkmale wird aber eine proportionale Stichprobenzie- hung benötigt, um die Repräsentativität zu gewährleisten.

Die beschriebene Stichprobenvariante er- möglicht zwar eine hinreichend genaue Feststellung der amtlichen Einwohnerzahl in allen Gemeinden, dennoch macht der Nachweis von Zusatzmerkmalen bis zur Kreisebene eine zusätzliche Erhebung in Gemeinden unter 10 000 Einwohnern erfor- derlich, weil gerade für Flächenländer mit einem hohen Anteil kleiner Gemeinden der Erfassungsgrad für Auswertungen unterhalb der Kreisebene ansonsten zu gering ist. Die Zahl der dort einzubeziehenden Adressen soll in dem beschriebenen Modell dem 500-Fachen des Bevölkerungsanteils der Gemeinden unter 10 000 Einwohner an der jeweiligen Kreisbevölkerung entsprechen. Dieser Stichprobenumfang gestattet, wenn überhaupt, lediglich einen Ergebnismachweis bis zur Kreisebene; ein flächendeckender Nachweis der Merkmale in tiefer regionaler

K 1

Zensusstichprobe (Modellvariante 2.2)



Gliederung ist nicht möglich. Ein Blick auf die Gemeindestruktur verschiedener Bundesländer (siehe Tabelle 8) zeigt, dass dies gerade für Rheinland-Pfalz ein nicht zu unterschätzendes Problem darstellt.

Werden Disproportionalität bei der Stichprobe außen vor gelassen, würden im Ergebnis bundesweit statt der angestrebten 8% bis 10% nur knapp 5% der Bevölkerung berücksichtigt. Die Stichprobe umfasst beispielsweise in den Stadtstaaten weniger Personen als dort in den Mikrozensus einbezogen werden.

Unzureichender Erfassungsgrad für kleinräumliche Nachweise unterhalb der Kreisebene ...

Auch für Flächenländer – insbesondere für Rheinland-Pfalz – ergibt sich mit rund 3% der Bevölkerung ein unzureichender Erfassungsgrad für kleinräumige Nachweise unterhalb der Kreisebene.

Mit einer vonseiten der statistischen Ämter vorgeschlagenen Erweiterung des Modells auf Gemeindeverbände und Stadtteile würde das Regionalisierungspotenzial insbesondere in kleinräumig strukturierten Gebieten und Großstädten deutlich erhöht. Dabei soll durch eine gezielte Berücksichtigung von Stadtteilen und Gemeindeverbänden der Stichprobenumfang erhöht werden.

... erfordert eine Erweiterung des Modells

Zusätzlich bedarf es einer Design-Modifikation, um annähernd den gleichen Informationsnutzen in der Fläche zu erzielen. Dabei sollen insbesondere regionale Besonderheiten eine stärkere Beachtung finden.

Mit dem Ziel, ein mit Blick auf die regionalen Datenanforderungen optimiertes Auswahl- und Schätzdesign zu entwickeln, wurde vom Statistischen Bundesamt ein Gutachten an die Universität Trier sowie an GESIS-ZUMA in Auftrag gegeben.

Von der erfolgreichen Umsetzung der vorgeschlagenen Modellerweiterungen bzw. -modifikationen wird es abhängen, inwieweit das Regionalisierungspotenzial des Zensus optimal ausgeschöpft werden kann. Jedenfalls wäre es ein Rückschlag für die amtliche Statistik, wenn die Regionalisierungspotenziale des registergestützten Zensus (deutlich) hinter denen einer traditionellen Vollerhebung zurückblieben. Mit Blick auf die Bedarfe der Landes-, Regional- und Kommunalplaner bleibt zu hoffen, dass die Bemühungen zur Entwicklung eines optimierten Auswahl- und Schätzdesigns von Erfolg gekrönt sein werden.

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

## Impressum

Überarbeitete Auszüge aus den Statistischen Monatsheften  
Rheinland-Pfalz 2/2009 bis 4/2009

Herausgeber:  
Statistisches Landesamt Rheinland-Pfalz  
Mainzer Straße 14-16  
56130 Bad Ems

Telefon: 02603 71-0  
Telefax: 02603 71-3150

E-Mail: [poststelle@statistik.rlp.de](mailto:poststelle@statistik.rlp.de)  
Internet: [www.statistik.rlp.de](http://www.statistik.rlp.de)

Redaktion: Romy Feldmann

Druck: Hessisches Statistisches Landesamt

Fotos: Jürgen Hammerl und Manfred Riege

---

© Statistisches Landesamt Rheinland-Pfalz  
Bad Ems · 2009

Für nichtgewerbliche Zwecke sind Vervielfältigung und unentgeltliche Verbreitung, auch auszugsweise, mit Quellenangabe gestattet. Die Verbreitung, auch auszugsweise, über elektronische Systeme/Datenträger bedarf der vorherigen Zustimmung. Alle übrigen Rechte bleiben vorbehalten.