

Schriftenreihe der ASI – Arbeitsgemeinschaft
Sozialwissenschaftlicher Institute

Bettina Klumpe · Jette Schröder
Markus Zwick *Hrsg.*

Qualität bei zusammengeführten Daten

Befragungsdaten, administrative
Daten, neue digitale Daten:
miteinander besser?



Springer VS

Schriftenreihe der ASI – Arbeitsgemeinschaft Sozialwissenschaftlicher Institute

Reihe herausgegeben von

Frank Faulbaum, Duisburg, Deutschland

Stefanie Kley, Hamburg, Deutschland

Birgit Pfau-Effinger, Hamburg, Deutschland

Jürgen Schupp, Berlin, Deutschland

Jette Schröder, Mannheim, Deutschland

Christof Wolf, Mannheim, Deutschland

Herausgegeben von

Frank Faulbaum
Universität Duisburg-Essen

Stefanie Kley
Universität Hamburg

Birgit Pfau-Effinger
Universität Hamburg

Jürgen Schupp
DIW Berlin

Jette Schröder
GESIS – Leibniz-Institut für
Sozialwissenschaften

Christof Wolf
GESIS – Leibniz-Institut für
Sozialwissenschaften

Weitere Bände in der Reihe <http://www.springer.com/series/11434>

Bettina Klumpe · Jette Schröder
Markus Zwick
(Hrsg.)

Qualität bei zusammengeführten Daten

Befragungsdaten, administrative
Daten, neue digitale Daten:
miteinander besser?

Hrsg.

Bettina Klumpe
ADM Arbeitskreis Deutscher Markt-
und Sozialforschungsinstitute e.V.
Berlin, Deutschland

Jette Schröder
GESIS – Leibniz-Institut für
Sozialwissenschaften
Mannheim, Deutschland

Markus Zwick
Statistisches Bundesamt
Wiesbaden, Deutschland

ISSN 2625-9427

ISSN 2625-9435 (electronic)

Schriftenreihe der ASI - Arbeitsgemeinschaft Sozialwissenschaftlicher Institute

ISBN 978-3-658-31008-0

ISBN 978-3-658-31009-7 (eBook)

<https://doi.org/10.1007/978-3-658-31009-7>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert durch Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2020

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Springer VS ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Inhaltsverzeichnis

Vorwort	3
<i>Frank Faulbaum</i>	
Einführende Bemerkungen	5
<i>Ralf Münnich</i>	
Qualität der regionalen Armutsmessung – vom Design zum Modell	11
<i>Sandra Hadam, Timo Schmid & Joanna Simm</i>	
Kleinräumige Prädiktion von Bevölkerungszahlen basierend auf Mobilfunkdaten aus Deutschland	31
<i>Barbara Wawrzyniak</i>	
Mehr Transparenz und Qualität mit innovativen Small Area Methoden	49
<i>Hartmut Scheffler</i>	
Datenzusammenführung und integrierte Analysen in Sozialforschung und Marktforschung: Erkenntnis- und Qualitätsgewinn	67
<i>Torsten Tümmler</i>	
Qualität bei zusammengeführten Daten. Versuch einer systematischen Annäherung	85
<i>Sebastian Götte & Bettina Klumpe</i>	
Die ADM Transparenz-Initiative	101
<i>Jan Goebel</i>	
Verbindung von Surveydaten und Geodaten. Möglichkeiten, Mehrwert und Probleme am Beispiel des SOEP	107

<i>Hanna Brenzel, Clara Schartner, Kathrin Gebers, Hannes Taubenböck & Michael Wurm</i>	
Geokoordinaten als Verknüpfungsmerkmal. Ein Werkstattbericht zur Zusammenführung von Daten der amtlichen Statistik mit Fernerkundungsdaten	123
<i>Hendrik Wagenseil</i>	
Vergleichbar, passgenau, ganzheitlich. Kunden-Mehrwert durch Datenintegration im Geomarketing	135
<i>Rainer Schnell</i>	
Record Linkage als zentraler Baustein der Forschung mit Registern und Big Data-Nutzungen	151
<i>Joseph W. Sakshaug</i>	
Linking Surveys with Big Data. Issues of Consent	167
Adressen der Referentinnen und Referenten	179

Vorwort

Die vorliegende Publikation dokumentiert die Beiträge der wissenschaftlichen Fachtagung „Qualität bei zusammengeführten Daten – Befragungsdaten, administrative Daten, neue digitale Daten: Miteinander besser?“, die am 27. und 28. Juni 2019 im Statistischen Bundesamt, Wiesbaden stattgefunden hat. Die Tagung ist die dreizehnte Veranstaltung einer Reihe wissenschaftlicher Fachtagungen, die das Statistische Bundesamt in Zusammenarbeit mit dem ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. und der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) seit dem Jahr 1995 in zwei-jährigen Abständen durchführt.

Die etablierte Veranstaltungsreihe gemeinsamer wissenschaftlicher Fachtagungen ist Themenbereichen gewidmet, die für Marktforscher/innen, Sozialwissenschaftler/innen und die amtliche Statistik gleichermaßen von Interesse sind. Sie fördert damit den intensiven, persönlichen Informations- und Erfahrungsaustausch zwischen den beteiligten Gruppen und trägt auf diese Weise zum wechselseitigen Verständnis der jeweiligen Forschungsinteressen und -herausforderungen bei.

Die inhaltliche Klammer der einzelnen Fachtagungen ist der Aspekt der Förderung und Sicherung der wissenschaftlichen Qualität der empirischen Forschung in akademischen und privatwirtschaftlichen Forschungseinrichtungen sowie bei der Datengewinnung in der amtlichen Statistik. In all diesen Institutionen werden zunehmend Daten aus verschiedenen Quellen zusammengeführt und gemeinsam analysiert – sei es um die Datenerhebung zu vereinfachen oder Analysen effizienter zu gestalten, sei es um neue Fragestellungen zu bearbeiten. Ziel der Tagung im Jahr 2019 war es daher, einen Einblick in aktuelle Anwendungsfälle der Zusammenführung von Daten zu geben sowie die mit der Zusammenführung verbundenen Chancen und Herausforderungen zu beleuchten.

Der Dank der Herausgeber gilt Frau Sibylle von Oppeln-Bronikowski und Herrn Prof. Dr. Frank Faulbaum für die Moderation der Veranstaltung, allen Referentinnen und Referenten für ihre Beiträge sowie Frau Bettina Zacharias und Herrn Marco Schwickerath für ihre engagierte Hilfe bei der Erstellung des Bandes. Nicht vergessen werden sollen darüber hinaus alle diejenigen, die durch ihre organisatorische und technische Unterstützung im Hintergrund die Durchführung der Tagung möglich gemacht und zu ihrem Gelingen beigetragen haben. Wir hof-

fen, dass auch dieser Band wie seine Vorgänger auf ein positives Echo stoßen wird und wünschen eine anregende Lektüre.

Berlin, Mannheim und Wiesbaden im November 2019

Bettina Klumpe

Jette Schröder

Markus Zwick



Einführende Bemerkungen

Frank Faulbaum

Vorstandsvorsitzender der ASI

1 Ziele der Datenzusammenführung

Die folgenden Ausführungen sollen die Ziele und wichtige Funktionen der Zusammenführung von Daten verdeutlichen. Der Begriff der Datenzusammenführung impliziert die Annahme, dass mehrere, möglicherweise aus unterschiedlichen Datenquellen stammende Datenmengen vorliegen, die Ergänzungen voneinander darstellen (vgl. Abbildung 1). In der Regel dürfte zunächst ein Datensatz vorliegen, der aus verschiedenen Gründen durch weitere Daten ergänzt wird.

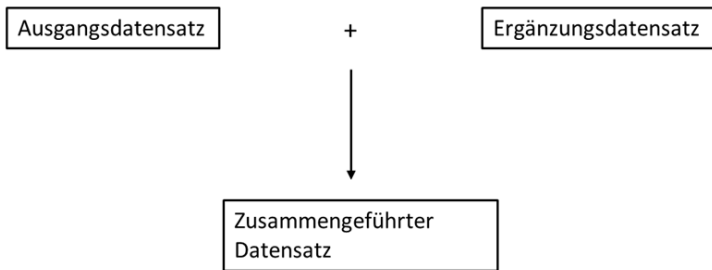


Abbildung 1 Ausgangsdatensatz und Ergänzungsdatensatz

Als Missing-Data-Problem betrachtet könnten die neu hinzu gekommenen Daten, wären sie noch nicht erhoben worden, in Bezug auf die Ausgangsdaten als fehlend aufgefasst werden. Die Zusammenführung wäre dann eine Ersetzung fehlender Daten (vgl. Goldstein und Harron 2015; Little und Rubin 2002). Die Verwendung des Begriffs „Datensatz“ sollte nicht darüber hinwegtäuschen, dass dieser nicht immer das Format von

Datensätzen in Statistik-Paketen haben muss. Grundsätzlich können die Daten nicht nur aus unterschiedlichen Quellen stammen, sondern auch ganz unterschiedlich organisiert sein. Eine statistische Analyse der zusammengeführten Daten erfordert dann möglicherweise einen mehr oder weniger großen Aufwand der Datenaufbereitung.

Die Zusammenführung von Daten, ob deterministisch oder probabilistisch, eröffnet einerseits die Möglichkeit von Analysen zusätzlicher Zusammenhänge und damit auch die Möglichkeit, neue statistische Modelle zu überprüfen. Andererseits können hinzugefügte Daten in ihrer Funktion als Hilfsvariablen zur Verbesserung der Schätzer von Populationsparametern beitragen, etwa als Hilfsvariablen für eine nachträgliche Schichtung. Die Zusammenführung von Daten aus unterschiedlichen Quellen und die Probleme ihrer methodischen und technischen Umsetzung unter Berücksichtigung der Privatheit, der De-Anonymisierungsgefahr und des Datenschutzes sind bereits seit einigen Jahren Thema der Umfrageforschung (vgl. die Beiträge in Harron, Goldstein und Dibben 2015). Ein wichtiges Anwendungsbeispiel ist die Verknüpfung von Umfragedaten mit administrativen Daten (vgl. z.B. Sakshaug et al. 2012; Schnell 2013). Da administrative Daten eine weitgehende Abdeckung der Population darstellen, ergibt sich in diesem Fall die Möglichkeit, die Beiträge abzuschätzen, die Coverage-Fehler und Nonresponse-Fehler zum totalen Umfragefehler (zum Begriff des totalen Umfragefehlers vgl. Faulbaum 2018) beitragen.

Neue Impulse für die Zusammenführung von Daten ergeben sich aus den Daten neuer Datenquellen, welche durch technologische Innovationen bereitgestellt werden können. Die technologischen Entwicklungen können dabei nie als abgeschlossen betrachtet werden. Beispiele für Daten neuer Quellen sind Sensordaten, die mit Hilfe von Einrichtungen in Smartphones erhoben werden können, Daten, die satellitengestützt erhoben werden wie GPS-Daten, Google-Earth-Daten, oder Daten, die durch Auswertungen von Reaktionen im Web erhoben werden (Auswertungen des Such- und Reaktionsverhaltens sowie von individuellen Handlungen wie z.B. Transaktionen im mobilen Web). Die Nutzung des Internets der Dinge befindet sich noch in der Startphase.

Die Erhebungsverfahren zahlreicher Arten neuer digitaler Daten sind sehr oft nicht-reaktiv. Einige Begriffe der traditionellen Klassifikation nicht-reaktiver Verfahren in Psychologie und Sozialwissenschaft lassen sich ohne Probleme mit leichten Änderungen in der Interpretation auf die neuen digitalen Daten übertragen. Physische Spuren entsprechen den Spuren, die unserer Handlungen im Netz hinterlassen, zur nicht-reaktiven Beobachtung könnte man die GPS-Ortung und die

unterschiedlichen Arten von Geodaten ebenso zählen wie die Erhebung von Sensordaten. Der Einsatz von Bots und die Beobachtung ihrer Wirkungen könnten als Feldexperimente organisiert werden und die Analyse laufender Berichte findet sich der Analyse von Webstatistiken wieder. Schließlich findet auch die Inhaltsanalyse ihre Anwendungen im Web, etwa bei der Analyse von Websites, Blogs oder von Kommunikationsformen wie Chats, WhatsApp-Mitteilungen etc. Auch die Analyse von Bildinformationen in Instagram wäre ebenfalls ein Beispiel.

Zusammenfassend kann die Zusammenführung von Daten folgende Funktionen erfüllen (vgl. Abbildung 1):

*Verbesserung der Schätzqualität der Zielvariablen
(Verbesserung der Anpassungsgewichtung):*

Diese Verwendung setzt voraus, dass sich die Variablen des Ergänzungsdatensatzes als Hilfsvariablen oder im erweiterten Sinne als Hilfsinformationen zur Optimierung der Schätzung der Zielvariablen des Ausgangsdatensatzes einsetzen lassen. Dies ist nur sinnvoll, wenn ein plausibler und vielleicht sogar statistisch evaluierbarer Zusammenhang zwischen den Hilfsvariablen des Ergänzungsdatensatzes und den Zielvariablen des Ausgangsdatensatzes angenommen werden kann. Wenn es keine plausible theoretische Verbindung gibt, ist die Funktion der Variablen als Hilfsvariablen eher fraglich. Allerdings könnte eine explorative Analyse Korrelationen identifizieren, die dann einer theoretischen Begründung unterzogen werden können.

Theoriegesteuerte Ergänzung des Ausgangsdatensatzes:

In diesem Fall steht die Erweiterung der Analysefragestellung im Mittelpunkt, etwa durch Bereitstellung zusätzlicher Prädiktoren wie etwa Kontextinformationen (Geoinformationen, Raster und/oder Funkzellinformationen, Informationen über den betrieblichen Kontext etc.) oder zu Einbeziehung Kontrollvariablen. Hier hinein fällt aber auch die Suche nach zusätzlichen Indikatoren für ein theoretisches Konstrukt.

Ergänzung des Ausgangsdatensatzes zur Absicherung der Reliabilität und Validität von Messungen:

Da die Reliabilität gleich der Korrelation zwischen parallelen Messungen ist (vgl. Lord und Novick 1968), könnten parallele Messungen im Ergänzungsdatensatz zu einer Abschätzung der Reliabilität dienen. Ein Beispiel wäre die satellitengesteuerte Abbildung der Bebauung einer Rasterzelle als Indikator der sozialen Schicht im Vergleich zu in

einer Umfrage erhobenen Schichtindikatoren. Allerdings erfordert die Nutzung von Satellitenbildern ein Rating der Bildinformation, die wiederum fehleranfällig ist und durch eine Interrater-Korrelation abgeschätzt werden sollte. Da die Quadratwurzel aus der Reliabilität gleich der theoretischen Validität ist, hätte man hiermit auch eine Schätzung der Validität.

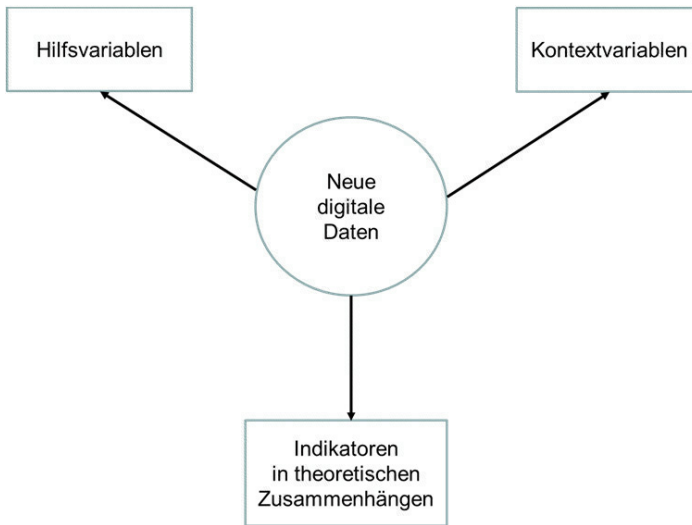


Abbildung 2 Funktionen neuer digitaler Daten

2 Ergänzende Daten als Indikatoren

Die „neuen digitalen Daten“, die z.T. in unübersehbarer Fülle als „big data“ anfallen, sind zunächst nur Daten und nichts sonst. Ihr Sinn ergibt sich erst entweder aus ihrer Funktion als Hilfsvariablen oder aus ihrer Integration in einen theoretischen Zusammenhang, in dem sie als Messungen von Inhalten angesehen werden können. Dies bedeutet wiederum, dass sie als Indikatoren für theoretische Variablen fungieren, deren Messeigenschaften wie Reliabilität und Validität hinterfragt werden sollten. Abbildung 3 zeigt beispielhaft ein Messmodell für eine theoretische Variable, im Beispiel die Variable „Lebensqualität“, die u.a. über einen Indikator „Lebenszufriedenheit“ gemessen wird. Aller-

dings ist der Indikator zunächst eine latente Antwortvariable, die als Ergebnis einer Messoperation noch in eine beobachtete empirische Variable überführt werden muss (vgl. Muthén 1984).

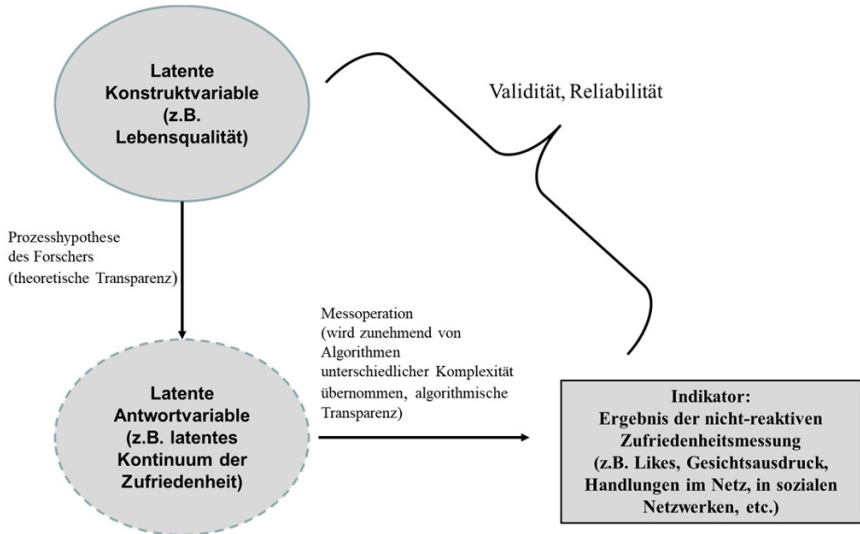


Abbildung 3 Messmodell mit „neuen digitalen Daten“

Als Messungen der latenten Antwortvariablen kommen nunmehr neue, auf neuen digitalen Daten beruhende empirische Variablen dazu, wie „Likes“, Ergebnisse von mathematischen Analysen des Gesichtsausdrucks oder auch Geodaten, welche die Attraktivität von Wohngebieten zum Ausdruck bringen. Die funktionalen Zusammenhänge im Modell sollten sowohl in der theoretischen Analyse als auch in den Messverfahren transparent sein. So sollten die Algorithmen zur Gesichtserkennung so transparent wie möglich sein.

Abbildung 3 zeigt die Grundstruktur eines Zusammenhangs zwischen theoretischen und empirischen Variablen, die auch bei Daten neuer Datenquellen nicht aus dem Auge verloren werden sollte. Dabei kann sich ein solches Modell auf ganz unterschiedliche, z.T. auf verschiedenen Niveaus aggregierte Daten beziehen. Hat man Messmodelle für die integrative Nutzung von neuen digitalen Daten formuliert, so lassen sich diese in komplexere theoretische Strukturen integrieren, die bei geeigneter methodischer Umsetzung statistisch überprüfbar sind.

3 Schlussbemerkung

„Neue digitale Daten“ haben u.a. die Funktion von *Messungen*, die den bekannten Qualitätskriterien genügen sollten, welche die Qualität einer Messung belegen können, auch wenn die dazu notwendigen statistischen Analysen nicht leicht vorzunehmen sind. Es scheint aber notwendig, auch für diese Daten praktikable methodische Vorschläge und Ideen dafür zu entwickeln, wie Kriterien der Reliabilität und Validität überprüft werden können. Die oft vorgetragene Behauptung der Objektivität der Daten reicht als Kriterium nicht aus.

Literatur

- Faulbaum, F. (2019). Total survey error. In J. Blasius & N. Baur (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (2. Auflage) (S. 505-519). Wiesbaden: Springer VS.
- Goldstein, H. & Harron, K. (2015). Record linkage: A missing data problem. In K. Harron, H. Goldstein, & C. Dibben (Eds.). *Methodological developments in data linkage* (S.109-124). Chichester, UK: John Wiley.
- Harron, K., Goldstein, H., & Dibben, C. (Eds.) (2015). *Methodological developments in data linkage*. Chichester, UK: John Wiley.
- Little, R.A. & Rubin, D.B. (2002). *Statistical analysis of missing data* (2nd edition). Hoboken, NJ: John Wiley.
- Lord, F.M. & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Muthén, B.O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49, 115-132.
- Sakshaug, J.W. et al. (2012). Linking survey and administrative records. *Sociological Methods & Research*, 41, 535-569.
- Schnell, R. (2013). *Linking surveys and administrative data*. German RCL Working Paper No. wp-grcl-2013-03.
- Schnell, R. (2015). Privacy preserving record linkage. In K. Harron, H. Goldstein, & C. Dibben (Eds.). *Methodological developments in data linkage* (S. 201-225). Chichester, UK: John Wiley.



Qualität der regionalen Armutsmessung – vom Design zum Modell

Ralf Münnich

Universität Trier

1 Zur Bedeutung der Qualität in Erhebungen

Mit dem European Statistics Code of Practice (European Statistical System, 2017) hat das Europäische Statistische System einen Verhaltenskodex verabschiedet, der Standards für Statistiken definiert. Gerade im Zeitalter von *Fake News* und *schnellen Statistiken* sind einheitliche und nachvollziehbare Standards in den Statistiken unverzichtbar. Selbstverständlich müssen diese Standards geeignet und vor allem nachvollziehbar umgesetzt werden.

Der Artikel befasst sich mit Methoden der Qualitätsmessung am Beispiel der Armutsmessung in Deutschland. Dabei spielen verschiedene aktuelle Diskussionen eine zentrale Rolle. Zum einen wird immer noch der Begriff der Repräsentativität in ungeeigneter und unreflektierter Form verwendet. Zum anderen werden immer mehr regionalisierte Statistiken von Staat und Gesellschaft verlangt. Dieser Bedarf liefert automatisch die Notwendigkeit, hochkomplizierte Methoden zu verwenden, und widerspricht oft auch der realen Umsetzbarkeit, welche durch Geheimhaltungsregeln konterkariert wird. Schließlich eröffnen neue digitale Daten und Webportale ungeahnte Möglichkeiten – zu meist unter Auslassung jedweder Angabe zur Qualität der so gewonnenen Statistiken.

1.1 Repräsentativität und Qualitätsmessung

Der Begriff der Repräsentativität ist seit jeher umstritten (eine aktuelle Übersicht kann Gabler und Häder, 2019, entnommen werden). Krug, Nourney und Schmidt (2001, Seite 18) definieren den Begriff wie folgt:

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert durch Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2020
B. Klumpe et al. (Hrsg.), *Qualität bei zusammengeführten Daten*, Schriftenreihe der ASI – Arbeitsgemeinschaft Sozialwissenschaftlicher Institute, https://doi.org/10.1007/978-3-658-31009-7_2

Eine Stichprobe gilt als repräsentativ, wenn sie statistisch gesicherte Rückschlüsse auf die Gesamtheit ermöglicht.

Hiermit werden bereits einige Voraussetzungen an den Begriff implizit verwendet. Zunächst geht man von einer Zufallsstichprobe aus (siehe auch Kauermann und Küchenhoff, 2011, Seite 10). Weiterhin impliziert ist, dass alle Strukturen der Grundgesamtheit in der Stichprobe enthalten sind. Vielfach wird in diesem Zusammenhang von *einem kleinen Abbild* der Grundgesamtheit gesprochen. Dies lässt sich jedoch ausschließlich im Sinne eines Zufallsmechanismus präzise definieren.

Nähert man sich dem Begriff der Repräsentativität mit Suchmaschinen, so findet man schnell, dass die Proportionen in der Grundgesamtheit sich in der Stichprobe wiederfinden müssen. Hier spiegelt sich die Idee wider, dass eine Stichprobe proportional in Bezug auf wichtige interessierende Merkmale ist. Zum einen stellt sich die Frage, welche Merkmale tatsächlich in diesem Zusammenhang als wichtig eingestuft werden sollten, zum anderen werden in Anwendungen in den seltensten Fällen proportionale Stichproben verwendet. Gerade bei Telefonstichproben sind a priori vergleichsweise wenig Details über die Gesamtheit bekannt, so dass diese Formulierung nicht zielführend wäre. Über diese einfachen Formulierungen hinaus, wird in der Wahrnehmung eine *repräsentative Stichprobe* oft automatisch als qualitativ hochwertig eingestuft, auch wenn keinerlei Informationen über die Genauigkeit im Sinne von Stichprobenfehlern oder die Präsenz und Auswirkungen von Nichtstichprobenfehlern, insbesondere in Folge von Nonresponse oder Rahmenfehlern, verfügbar sind.

Schnell (2019, Seite 163) weist darauf hin, dass bei repräsentativen Stichproben drei technische Fragen positiv beantwortet werden müssen. Neben dem Vorhandensein einer vollständigen Liste der Erhebungseinheiten¹ müssen die Inklusionswahrscheinlichkeiten für alle Elemente positiv und berechenbar sein. Auf Basis dieser Formulierung lassen sich statistische Eigenschaften von Schätzverfahren, insbesondere Unverzerrtheit und Effizienz, konkretisieren. Darüber hinaus, und das ist gerade bei Personen- und Haushaltsbefragungen von besonderer Bedeutung, spielt der Mechanismus des Antwortausfalls sowie dessen Ausmaß eine ganz erhebliche Rolle. Gerade in Zeiten, in denen die Response-Quoten immer geringer werden, wird die Frage bedeutsa-

1 An dieser Stelle sei angemerkt, dass Methoden des Indirect Sampling keine vollständige Liste von Erhebungseinheiten benötigen (Lavalée, 2009).

mer, ob und inwieweit Verfahren der Korrektur von Antwortausfällen noch zu sinnvollen Ergebnissen führen.

Gabler und Quatember (2013) geben einen Überblick über die Entwicklung und die aktuelle Diskussion zum Begriff der Repräsentativität. Hierbei weisen sie sowohl auf Stichproben- wie auch Nichtstichprobenfehler hin, welche bei der Umsetzung einer repräsentativen Stichprobe beachtet werden müssen. Ganz zentral in ihrer Darstellung ist die Formulierung eines Mindeststichprobenumfangs, um bestimmte Qualitätsbedingungen für die zu Grunde liegende Schätzstatistik zu erfüllen. Sie gehen hierbei zunächst von (approximativer bzw. asymptotischer) Unverzerrtheit der interessierenden Schätzstatistik aus. Anschließend werden derartige Betrachtungen stets auf Varianzschätzungen bzw. Funktionen dieser, wie etwa auf den relativen Standardfehler oder die Konfidenzintervalllänge der Schätzstatistik, zurückgeführt. Darüber hinaus gehen sie auf die Bedeutung des Stichprobendesigns für die Varianzschätzung ein, bzw. wie diese mit Hilfe des Design-Effekts formuliert wird. Da heutzutage selten nur Schätzwerte für Gesamtheiten benötigt werden, sondern vielmehr auch für zahlreiche Subgruppen, erweitern sie die Betrachtungen auf die Repräsentativität bei Subgruppen.

Eine eingehende Darstellung dieser Methoden findet sich etwa in der einschlägigen Stichprobenliteratur (siehe z.B. Särndal, Swensson und Wretman, 1992, oder Lohr, 2010) oder in spezialisierten Werken wie Wolter (2007) oder Shao und Tu (1995). Ebenso muss der Abschlussbericht der Task Force Variance Estimation (siehe European Commission, 2002) bzw. das DACSEIS-Projekt (<http://www.dacseis.de>) genannt werden. Möchte man eine Statistik mit vorgegebener Mindestqualität erstellen, werden oft zwei Konzepte bemüht, welche aber prinzipiell äquivalent sind. Einerseits kann der maximale Standardfehler $SE(\mathcal{G})$ einer Statistik \mathcal{G} vorgegeben werden. Dies führt dann bei einer Totalwertschätzung für \mathcal{G} unter Verwendung des einfachen Urnenschemas zu

$$\hat{V}(\mathcal{G}) = N^2 \cdot \frac{\sigma^2}{n} \leq V_{\max}.$$

Hieraus kann dann ein Mindeststichprobenumfang n ermittelt werden. Andererseits kann auch eine Berechnung des minimalen Stichprobenumfangs über eine vorgegebene maximale Konfidenzintervalllänge definiert werden. Hier kommt lediglich ein weiterer zu berücksichtigender Korrekturfaktor hinzu. Die Formeln für komplexe Stichproben-