

Zur Datenqualität der Bildungsangaben im Mikrozensus

Schimpl-Neimanns, Bernhard

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Schimpl-Neimanns, B. (2006). *Zur Datenqualität der Bildungsangaben im Mikrozensus*. (ZUMA-Arbeitsbericht, 2006/03). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-200565>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

ZUMA-Arbeitsbericht Nr. 2006/03

**Zur Datenqualität der Bildungsangaben
im Mikrozensus**

Bernhard Schimpl-Neimanns

Mai 2006

ISSN 1437-4110

ZUMA
Quadrat B2, 1
Postfach 12 21 55
68072 Mannheim
Telefon: +49 621-1246-263
Telefax: +49 621-1246-100
E-mail: schimpl-neimanns@zuma-mannheim.de

Zusammenfassung

In diesem Beitrag wird am Beispiel des Besuchs der Klassenstufen 11-13 allgemein bildender Schulen die Qualität der Bildungsangaben im Mikrozensus 1996 diskutiert. Ergänzend zum Schulbesuch werden auf Basis des Mikrozensuspanels erstmals Analysen zur Antwortkonsistenz der Angaben zum allgemeinen Schulabschluss vorgestellt. Vergleiche mit den Populationsdaten aus der amtlichen Bildungsstatistik zum Schulbesuch lassen Erhebungs- und Abgrenzungsprobleme im Mikrozensus erkennen. Teilweise sind Schüler beruflicher Schulen entgegen den Definitionen des Mikrozensus als Besucher allgemein bildender Schulen erfasst. Zudem ist eine gravierende Übererfassung bei den unter 18-jährigen Oberstufenschülern festzustellen, die auf eine problematische Unterscheidung der Klassenstufen 5-10 vs. 11-13 bzw. der Sekundarstufen I und II verweist. Die mit dem Mikrozensuspanel berechneten Übergangsraten der Bildungsabschlüsse zwischen verschiedenen Zeitpunkten weisen zumeist eine Stabilität von über 80 Prozent auf. Im Vergleich zu Ergebnissen sozialwissenschaftlicher Umfragen spricht dies für eine ausreichende bis gute Datenqualität des allgemeinen Schulabschlusses.

Abstract

By the example of academic secondary school attendance the paper addresses data quality issues in the reporting of education data in the German Microcensus 1996. In addition to school attendance, new analyses on item stability of data on general educational certificates based on the German Microcensus-Panel are presented. A comparison of the Microcensus data with data from official Education Statistics points to certain limitations in data collection and definitions used in the Microcensus. It is not possible to distinguish adequately between students or graduates from general education schools and those from vocational schools. The Microcensus estimates a much higher enrollment of upper secondary level pupils aged under 18 than the Education Statistic. This suggests difficulties of respondents to differentiate lower from upper secondary levels (grades 5-10 vs. 11-13). Transition rates of the general educational attainment based on Microcensus-Panel data between different time points show a stability of over 80 per cent for the most categories. Compared with results of science based surveys this indicates sufficient to good data quality.

1. Einleitung¹

Für die empirische Sozial- und Wirtschaftsforschung zählt der Mikrozensus zu den wichtigsten amtlichen Datenquellen für bildungsstatistische Analysen. Im Vergleich zur amtlichen Bildungsstatistik, die sich in die Schul-, Berufsschul- und Hochschulstatistik gliedert, lassen sich eine Reihe von Fragestellungen nur mit dem Mikrozensus bearbeiten. Dies betrifft insbesondere alle Sachverhalte, bei denen es um den Haushalts- und Familienkontext geht. Während beispielsweise in der Schulstatistik Angaben zur sozioökonomischen Lage der Eltern völlig fehlen, können diese Informationen für die noch bei ihren Eltern lebenden Schüler im Mikrozensus verwendet werden. Des Weiteren können mit dem Mikrozensus Aspekte der Verwertung von Bildungsqualifikationen auf dem Arbeitsmarkt oder Zusammenhänge zwischen Bildungsstatus und Lebenslage untersucht werden.

Für die Beurteilung von Analyseergebnissen des Mikrozensus sind, wie bei allen anderen Bevölkerungsumfragen, Informationen über die Datenqualität unabdingbar. Während zu den erwerbsstatistischen Themen des Mikrozensus solche Qualitätsuntersuchungen ansatzweise vorliegen (Dräther et al. 2001; Pöschl 1992; Rudolph 1998; Schupp et al. 1999), sind zu den bildungsstatistischen Angaben des Mikrozensus bislang kaum Einschätzungen möglich. Zu den wenigen Ausnahmen zählen insbesondere Untersuchungen zu den Auswirkungen der Freiwilligkeit der Auskünfte auf die Qualität der Bildungsangaben (Esser et al. 1989; Riede und Emmerling 1994), in Bezug auf den Hochschulabschluss frühere Vergleiche des Mikrozensus mit der Hochschulstatistik (Esser et al. 1989: 131) sowie zur beruflichen Weiterbildung ein Vergleich mit dem Berichtssystem Weiterbildung (Lois 2005).

Verteilungen des Mikrozensus dienen aufgrund des Stichprobenumfangs und der Auskunftspflicht für viele andere Stichproben in und außerhalb der amtlichen Statistik als Hochrechnungsrahmen bzw. werden für Prüfungen selektiver Befragungsausfälle, die häufig eng mit der Bildungsqualifikation der ausgewählten und zu befragenden Personen verbunden sind, herangezogen. Vor diesem Hintergrund der Verwendung des Mikrozensus als Referenzstatistik stellen sich deshalb besondere Anforderungen an die Datenqualität.

Für die Qualitätsbeurteilung liefern Panelangaben wichtige Informationen insofern auf die Angaben aus vorherigen Befragungen zurückgegriffen werden kann. Damit werden unter anderem Fragen zur Antwortkonsistenz bearbeitbar. Seit dem Mikrozensusgesetz 1996 dürfen

¹ Dieser Aufsatz ist die schriftliche Ausarbeitung eines Vortrages bei der gemeinsamen Tagung der Arbeitsgemeinschaft sozialwissenschaftlicher Institute und der Methodensektion der DGS „Stichprobenqualität in Bevölkerungsumfragen“ am 14.-15. Oktober 2005 in Berlin. Für wertvolle Hinweise zu dem hier behandelten Thema danke ich Thomas Riede, Michael Blohm, Rolf Porst und Julia Schroedter danke ich für konstruktive Anmerkungen zu früheren Fassungen.

die statistischen Ämter die für die Zusammenführung der Querschnittsangaben zu einem Panel benötigten Ordnungsnummern speichern. In einem vom Bundesministerium für Bildung und Forschung (BMBF) und der Deutschen Forschungsgemeinschaft (DFG) finanzierten Projekt wurden die Voraussetzungen für die Weitergabe des Mikrozensus als Panel an die Wissenschaft geschaffen.² Für die Einschätzung des Analysepotenzials wurden in diesem Projekt unter anderem Bildungsverläufe untersucht. Daraus werden im Folgenden einige Befunde zur Datenqualität der Angaben zum Besuch der gymnasialen Oberstufe aufgegriffen (Schimpl-Neimanns 2005). Ergänzend zum Schulbesuch wird die zeitliche Konsistenz der Angaben zum allgemeinen Schulabschluss berichtet.

In den bisherigen Untersuchungen zum Besuch der gymnasialen Oberstufe wurde deutlich, dass Schüler bzw. Absolventen allgemein bildender Schulen und beruflicher Schulen nicht den Definitionen entsprechend unterscheidbar sind. Um den vermuteten Abweichungsgründen nachzugehen, werden in diesem Aufsatz zusätzliche Vergleiche mit der Bildungsstatistik vorgenommen. Seit 1991 wird der Schulbesuch in Anlehnung an die Internationale Standardklassifikation für das Bildungswesen (ISCED) im Mikrozensus nach Klassenstufen erhoben. Da der Schulbesuch bis 1990 nach Schularten erfasst wurde, kann insbesondere der Vergleich von Ergebnissen des Mikrozensus 1989 mit der Bildungsstatistik darüber Aufschluss geben, ob und in welcher Weise die bei den Mikrozensen 1996 bis 1999 gefundenen Fehlklassifikationen auf im Alltag eher schwierige Begriffe zurückzuführen sind.

Im Folgenden werden zunächst die verwendeten Daten und die Methode beschrieben sowie Gründe für die bei dem Vergleich der Datenquellen auftretenden möglichen Abweichungen genannt. Im vierten Abschnitt werden Verteilungen zum Besuch der gymnasialen Oberstufe im Mikrozensus 1996, 1991 und 1989 den Ergebnissen der Bildungsstatistik gegenüber gestellt. Anschließend werden erste Ergebnisse des Mikrozensuspanels 1996-1999 zur Antwortstabilität des allgemeinen Schulabschlusses berichtet. Der Aufsatz schließt mit einer zusammenfassenden Einschätzung zu den Verteilungsabweichungen.

2. Datenbeschreibung

Seit 2005 wird der Mikrozensus als unterjährige Erhebung durchgeführt. Mit dieser Umstellung auf eine kontinuierliche Befragung sind weitere Modifikationen des Frageprogramms sowie methodische Änderungen verbunden (siehe Afentakis und Bihler 2005; Lotze und Breiholz 2002a,b). Die folgende Darstellung bezieht sich auf die verwendeten Daten des Mikrozensus 1989 bis 1996 mit Konzentration auf das von 1990 bis 2004 geltende Stichpro-

² Siehe dazu die WWW-Seite des Projektes www.destatis.de/mv/mzpanel_start.htm.

bendesign und Frageprogramm. Im Anschluss an die Kurzbeschreibung der Bildungsstatistik wird die Konstruktion vergleichbarer Bildungsdaten skizziert.

2.1 Mikrozensus

Im Mikrozensus werden seit 1957 in Westdeutschland und seit 1991 in den neuen Bundesländern jährlich vielfältige Informationen über die demografische, soziale und wirtschaftliche Struktur der Bevölkerung erhoben. Die Stichprobe umfasst ein Prozent der Personen und Haushalte. Aufgrund der Auskunftspflicht liegt die Teilnahmequote der Haushalte bei rund 97 Prozent. In Bezug auf das Stichprobendesign ist der Mikrozensus als mehrfach geschichtete einstufige Klumpen- bzw. Flächenstichprobe und als Rotationspanel gekennzeichnet (Meyer 1994). Die Primäreinheiten in der Stichprobe bilden über 40.000 Auswahlbezirke (Klumpen). Sie bestehen in den Mikrozensus ab 1990 i. d. R. aus benachbarten Wohnungen, die auf Basis von Ergebnissen der Volkszählung 1987 bzw. des Zentralen Einwohnerregisters in den neuen Bundesländern gebildet wurden. Die Auswahlbezirke umfassen durchschnittlich etwa neun Wohnungen; im Mikrozensus bis 1989 waren es noch rund 23 Wohnungen. Die jährliche Aktualisierung der Mikrozensus-Stichprobe berücksichtigt die neu entstandenen Wohnungen bzw. Bebauung ganzer Flächen durch Verwendung von Daten der Bautätigkeitsstatistik.

Der Mikrozensus ist als eine Panelstichprobe angelegt, bei der die Haushalte eines Auswahlbezirkes vier Jahre lang befragt werden, wobei jedes Jahr ein Viertel der Auswahlbezirke ausgetauscht wird. Wegziehende Personen und Haushalte werden jedoch nicht weiter befragt, sondern durch die nachziehenden Personen bzw. Haushalte ersetzt (Prinzip der Flächenstichprobe).

In Bezug auf die Befragungsmethode dominiert die persönliche Befragung. Rund 15 Prozent der Interviews basieren auf schriftlichen Auskünften der Befragten. Fremdauskünfte (Proxy-Interviews) sind möglich, z. B. wenn Eltern die Fragen zum Schulbesuch und Bildungsabschluss ihrer Kinder beantworten. Erst ab der Erhebung 1999 wird im Rahmen der Unterstichprobe des Mikrozensus erfasst, ob es sich um Selbst- oder Fremdauskünfte handelt. Insgesamt beruhen knapp 30 Prozent der Angaben auf Proxy-Interviews (Breiholz 2000). Bis zum Mikrozensus 2004 bezieht sich der Berichtszeitraum i. d. R. auf die letzte feiertagsfreie Woche im April eines Jahres (Berichtswoche).

Die meisten Fragen sind auskunftspflichtig. Daneben gibt es weitere Fragen und Themenbereiche, die von der Auskunftspflicht ausgenommen sind. Hierzu zählen beispielsweise Fragen der EU-Arbeitskräftestichprobe (European Labour Force Survey), die als Unterstichprobe in

den Mikrozensus integriert ist, Fragen zur Gesundheit sowie (seit 1996) die Angaben zum allgemeinen und beruflichen Abschluss von Personen über 50 Jahren.

Die Fallzahlen des Mikrozensus werden in einem zweistufigen Verfahren hochgerechnet. Auf der ersten Stufe erfolgt eine Korrektur des Unit-Nonresponse von rund drei Prozent für die nicht erreichten Haushalte. Auf der zweiten Stufe werden die nach dieser Ausfallkorrektur gewichteten Mikrozensusergebnisse an bekannte demografische Randverteilungen angepasst. Diese Anpassung an Ergebnisse der laufenden Bevölkerungsfortschreibung erfolgt für die Merkmale Geschlecht und Staatsangehörigkeit (Deutsche vs. Ausländer) auf der regionalen Ebene von so genannten Anpassungsschichten sowie für Soldaten und Wehrpflichtige an entsprechende Bestandsmeldungen (siehe Heidenreich 1994). Die Anpassungsschichten umfassen durchschnittlich wenigstens 500.000 Einwohner.

Für die folgenden Auswertungen werden Daten der Mikrozensus Scientific Use Files verwendet. Dabei handelt es sich jeweils um faktisch anonymisierte 70-Prozent-Haushaltssubstichproben des vollen Mikrozensus. Das Ziehungsverfahren berücksichtigt die wesentlichen Designelemente des Mikrozensus und gewährleistet eine sehr hohe Übereinstimmung mit Ergebnissen der Originaldaten (Rendtel und Schimpl-Neimanns 2001). Das verwendete Panelfile des Mikrozensus umfasst die im Zeitraum von 1996 bis 1999 befragten Personen (Rotationsviertel). Die hier genutzten faktisch anonymisierten Paneldaten entsprechen zirka einer 65-Prozent-Haushaltssubstichprobe.

2.2 Bildungsstatistik

Die Statistik zu allgemein bildenden Schulen basiert auf einer Vollerhebung und ist als Teil der amtlichen Bildungsstatistik in den einzelnen Bundesländern unterschiedlich organisiert. Durch die Orientierung an einem Minimalprogramm können aber Bundesergebnisse zusammengestellt werden (Statistisches Bundesamt 1996; Weishaupt und Fickermann 2001). Die Statistik beruht auf Meldungen der Schulleiter (Anstaltsbefragung). Räumlichen Darstellungen, wie beispielsweise dem Schulbesuch nach Bundesland, liegt der Schulort zugrunde. Die summarischen Angaben umfassen unter anderem die Zahl der Schüler nach Klassenstufe, Alter, Geschlecht und Staatsangehörigkeit. Dass Klassenstufe und Alter nicht in Kombination ausgewiesen werden, schränkt Vergleiche mit dem Mikrozensus ein. Zu beachten ist auch, dass zur Verringerung der Belastung der Auskunft gebenden Schulen der Schulbesuch nicht in jedem Jahr und nicht in jedem Bundesland differenziert erhoben wird. Beispielsweise werden die nach Geburtsjahr gegliederten Schülerzahlen teilweise nur im Abstand mehrerer

Jahre ermittelt. Die Altersverteilungen der dazwischen liegenden Zeitpunkte werden auf Basis der ausführlichen Strukturhebungen geschätzt.

Bei der Schulstatistik handelt es sich um Stichtagerhebungen, deren Berichtszeitraum sich jeweils auf den Beginn eines Schuljahres (Herbst) bezieht. Die Angaben über Schulklassene nach Abschlussart beziehen sich auf das vergangene Schuljahr bzw. das Schuljahresende.

2.3 Vergleichbare Erfassung der gymnasialen Oberstufe

In den alten Bundesländern und Brandenburg bilden bis Ende der 1990-er Jahre die Klassenstufen 11-13 allgemein bildender Schulen der Sekundarstufe II die gymnasiale Oberstufe. In Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt und Thüringen kann das Abitur schon nach 12 Schuljahren absolviert werden. Der klassische Abschluss der gymnasialen Oberstufe ist die allgemeine Hochschulreife (Abitur) oder die fachgebundene Hochschulreife. In einzelnen Bundesländern wird bei vorliegendem Versetzungszeugnis in die 13. (bzw. 12.) Klasse der Abschluss der Fachhochschulreife vergeben, die eigentlich ein beruflicher Bildungsabschluss ist. Neben der Fachhochschulreife können sowohl die fachgebundene Hochschulreife als auch die allgemeine Hochschulreife auch an beruflichen Schulen erlangt werden (KMK 2003). In den meisten Bundesländern zählen zur gymnasialen Oberstufe Gymnasien, Integrierte Gesamtschulen und Freie Waldorfschulen, daneben die Einrichtungen Kolleg und Abendgymnasium des zweiten Bildungsweges. Die gymnasiale Oberstufe ist in den einzelnen Bundesländern unterschiedlich organisiert (Bellenberg et al. 2004). Während die Kultusministerkonferenz in ihrer Übersicht des deutschen Schulsystems bei der Abgrenzung der gymnasialen Oberstufe eine breite Definition verwendet (KMK 2001), die auch berufliche Gymnasien einschließt, wird sowohl im Mikrozensus als auch in der Schulstatistik strikt zwischen allgemein bildenden und beruflichen Schulen unterschieden. Berufliche Gymnasien mit dem Bildungsziel Abitur oder Fachhochschulreife (Fachoberschulen, Fachgymnasien, Berufs-/Technische Oberschulen) gehören demnach zu den beruflichen Schulen.

Seit 1991 wird im Mikrozensus die Filterfrage nach dem gegenwärtigen Schulbesuch gestellt. Schulbesucher werden nach der Art der Schule gefragt, wobei Schüler einer allgemein bildenden Schule die Klassenstufe (1-4, 5-10, 11-13) angeben können. Die Kategorie „Klassenstufe 11-13“ ist in Klammern mit dem Zusatz „gymnasiale Oberstufe“ versehen. An dieser Stelle ist anzumerken, dass Befragte, die sich für eine schriftliche Beantwortung mittels Selbstausfüller-Fragebogen entschieden haben, im Fragebogen keine weiteren Erläuterungen zu den abge-

fragten Begriffen erhalten.³ Eine Gegenüberstellung der Schularten zur gymnasialen Oberstufe bzw. zu den beruflichen Schulen findet sich nur im Interviewerhandbuch. Da rund 85 Prozent aller Angaben auf klassischen Interviews beruhen, kommt es daher in erster Linie darauf an, ob der Interviewer die genauen Definitionen kennt und anwendet.

Bis 1990 wurde der Schulbesuch im Mikrozensus ohne die Unterscheidung allgemein bildender und anderer Schulen und nicht nach Klassenstufen erhoben, sondern es wurde nur die jeweilige Schulform erfragt. Dabei waren Schüler aller Einrichtungen mit dem Bildungsziel Fachhochschulreife und Hochschulreife in der Kategorie „Gymnasium/Fachoberschule“ anzugeben. Im Unterschied zu den Erhebungen ab 1991 sind damit 1989 berufliche Gymnasien den allgemein bildenden Schulen der Sekundarstufe II zugeordnet. Der Besuch der Gymnasialzüge an Gesamt- und Sonderschulen (Sekundarstufe II) war ebenfalls in der Kategorie „Gymnasium/Fachoberschule“ einzutragen. Die genauen Zuordnungen werden wiederum lediglich im Interviewerhandbuch beschrieben.

Die im Mikrozensus verwendeten Definitionen stimmen weitestgehend mit der Schulstatistik überein bzw. die jeweiligen Schulformen können vergleichbar abgegrenzt werden. Für den Vergleich des Mikrozensus mit der Schulstatistik ist dies wichtig, da somit konzeptionelle Abweichungen ausgeschlossen werden können.

3. Methode und Abweichungsgründe

Datenqualität bezieht sich allgemein darauf, in welcher Weise statistische Ergebnisse den mit der Datenerhebung verfolgten Zweck erfüllen. Zu den wichtigsten Aspekten zählen Schätzgenauigkeit und Reliabilität sowie Validität, aber auch die Verfügbarkeit und der Datenzugang (Office of Management and Budget 2001: 1-2). Häufig steht bei Fragen der Datenqualität die Genauigkeit bzw. der Stichprobenfehler im Vordergrund. Jedoch dürfen bei der Frage der Datenqualität systematische Fehler keinesfalls ausgeblendet werden, da sie in vielen Umfragen größer als der Stichprobenfehler sein können (Särndal et al. 1997: 539).

Die Aufdeckung systematischer Fehler ist nicht einfach und setzt oftmals externe Datenquellen voraus. Im Idealfall werden verschiedene Verfahren kombiniert. Dazu gehören Pretests zur Kontrolle des Fragebogens und Wiederholungsbefragungen mit anderen Frageversionen bzw. differenzierten Nachfragen zu einzelnen Sachverhalten, die zugleich der Interviewerkontrolle dienen können. In wenigen Fällen sind Verknüpfungen von Umfragedaten mit Re-

³ Die Antwortkategorien im Mikrozensus 1991 sind: Kindergarten/-krippe (0); Allgemeinbildende Schule, Klassenstufe 1-4 (1), Klassenstufe 5-10 (2), Klassenstufe 11-13 (gymnasiale Oberstufe) (3); Berufliche Schule (4); Fachhochschule (5); Hochschule (6). Die Fragebögen sind unter www.gesis.org/Dauerbeobachtung/GML/Daten/MZ/index.htm zugänglich.

gisterdaten und ein Abgleich der in beiden Datenquellen erfassten Merkmale möglich. Häufiger werden Randverteilungen mit anderen Daten verglichen. Bei den standardgemäß durchgeführten Analysen US-amerikanischer statistischer Ämter zur Datenqualität kommen diese Methoden häufig zum Einsatz (siehe unter anderem: Black et al. 2003; Brick et al. 1996; McGovern und Bushery 1999; NCES 1997; U.S. Census Bureau 2004).

Im Folgenden werden lediglich Randverteilungsvergleiche, also einfache deskriptive Vergleiche hochgerechneter Mikrozensusergebnisse mit der Bildungsstatistik durchgeführt. Dabei wird nicht vorausgesetzt, dass die Bildungsstatistik "wahre" Werte liefert. Allerdings wird angenommen, dass die Ergebnisse der Totalerhebung Bildungsstatistik weniger fehlerbehaftet sind als die des Mikrozensus. Diese Annahme erscheint gerechtfertigt.

Als Quellen systematischer Fehler kommen unklare Fragen oder Filterführungen im Fragebogen, Probleme der Erhebungs- bzw. Befragungsmethode, Interviewereinflüsse oder Verständnisschwierigkeiten und fehlerhafte Auskünfte sowie Antwortausfälle vonseiten der Befragten in Frage. Aus dem Vergleich mit der Bildungsstatistik kann nur die Gesamtabweichung ("Total Survey Error"; siehe Biemer und Lyberg 2003) ermittelt werden. Inwiefern die Gesamtabweichungen mit dem Stichprobenfehler, unterschiedlichen Berichtszeiträumen, Unit- und Item-Nonresponse, verschiedenen Antwortfehlern (u. a. aufgrund sozialer Erwünschtheit) oder unterschiedlichen Befragungsarten (Mode-Effekte) zusammenhängen, kann nicht geklärt werden. Zur ansatzweisen Kontrolle werden, sofern möglich, weitere Differenzierungen vorgenommen.

Die den Gesamtabweichungen zwischen Mikrozensus und Bildungsstatistik zugrunde liegenden Ursachen können wie folgt zusammengefasst werden.

- **Berichtszeitraum:** Während die Bildungsstatistik die Schülerzahlen zum Schulanfang eines Schuljahres (Herbst) berichtet, beziehen sich die Mikrozensusergebnisse auf die Berichtswoche des Folgejahres, die i. d. R. in der letzten feiertagsfreien Aprilwoche liegt. Die Unterschiede, die auf die Differenz zwischen Herbst des Vorjahres und April zurückzuführen sind, dürften jedoch gering sein, da Schulartwechsel und Ausbildungsabbrüche hauptsächlich zum Schuljahresende stattfinden.
- Zu berücksichtigen ist aber, dass im Mikrozensus nach dem "gegenwärtigen" Schulbesuch gefragt wird. Bei einem bis in den Sommer reichenden Befragungszeitraum kann die Fokussierung auf den Interviewzeitpunkt dazu führen, dass Schüler, die in den Schulferien, d. h. nach Abschluss des Schuljahres befragt werden, Angaben zum nächsten Schuljahr (Herbst) statt zum vergangenen Schuljahr machen. Gegebenenfalls wird auch der zwi-

schenzeitlich erreichte Schulabschluss statt der noch im April besuchten Schule angegeben. Grundsätzlich nicht auszuschließen ist auch der Fall, dass Angaben exakt zum Befragungszeitpunkt erfolgen, d. h. eventuell kein Schulbesuch berichtet wird. Allerdings ist damit kaum zu rechnen, sofern von den befragten Haushalten auch in den Schulferien der Hauptstatus eines Schülers wahrgenommen wird.

- **Merkmalsunterschiede bei Drittvariablen:** Drittvariablen liegen in der Bildungsstatistik nur sehr begrenzt vor. In dem Zusammenhang, dass in der Bildungsstatistik das Alter bzw. das Geburtsjahr nicht jedes Jahr direkt erhoben, sondern entsprechende Altersverteilungen teilweise geschätzt werden (s. o.), kann es zu Abweichungen zum Mikrozensus kommen. Diese dürften jedoch eher marginal sein.
- **Populationsschätzung, Hochrechnung:** Frühere Vergleiche zwischen Mikrozensus und Volkszählung haben gezeigt, dass die laufende Bevölkerungsfortschreibung, an die die Mikrozensusdaten angepasst werden, den Bevölkerungsbestand mit zunehmendem Abstand zur Volkszählung überschätzt (siehe zusammenfassend Rendtel und Schimpl-Neimanns 2001; Statistische Ämter des Bundes und der Länder 2004). Durch die Anpassung können somit Fehler in den Mikrozensus übertragen werden. Die Abweichungen zu den nicht extra angepassten Ergebnissen der Bildungsstatistik sollten sich bei Kontrolle weiterer Variablen, die wie beispielsweise das Alter nicht in der Anpassung berücksichtigt werden, jedoch in Form einer konstanten Verschiebung zeigen.
- **Definitionen im Mikrozensus:** Wie oben angesprochen, zählen im Mikrozensus ab 1991 ausschließlich Schüler allgemein bildender Schulen der Klassenstufen 11-13 zur gymnasialen Oberstufe. Schüler beruflicher Schulen mit den Bildungszielen Fachhochschulreife oder allgemeine Hochschulreife sind dagegen als Schüler beruflicher Schulen einzuordnen. Es ist davon auszugehen, dass diese Definition nicht dem Alltagsverständnis entspricht und deshalb in der Befragungspraxis nicht adäquat umgesetzt werden kann.
- **Befragungsart (Mode-Effekte):** Bei den Besuchern der gymnasialen Oberstufe liegt der Anteil der schriftlichen Auskünfte bei 14 Prozent. Da die genauen Definitionen nur im Interviewerhandbuch zu finden sind, nicht aber im Selbstausfüller-Fragebogen, ist damit zu rechnen, dass schriftlich Befragte bei der Beantwortung der Frage nach dem Schulbesuch auf ihre Alltagsdefinition zurückgreifen. Geht man aber davon aus, dass die Definitionen selbst im mündlichen Interview nicht immer umfassend berücksichtigt werden, dürften die unterschiedlichen Befragungsarten nur geringe Effekte besitzen.
- Bei Oberstufenschülern liegen im Mikrozensus 1999 mit 82 Prozent überdurchschnittlich häufig Fremdauskünfte vor. Bei Proxy-Interviews wird häufig eine eingeschränkte Daten-

qualität angenommen. Ob dies auch in Bezug auf die Angaben zum Schulbesuch zutrifft, ist fraglich, denn anders als bei Proxy-Angaben zum Einkommen oder zu Arbeitsstunden (Dawe und Knight 1997) kann man in diesem Fall davon ausgehen, dass Eltern über die von ihren Kindern besuchte Schule gut informiert sind.

4. Vergleiche zum Besuch der gymnasialen Oberstufe im Mikrozensus und in der Bildungsstatistik

Wie eingangs erwähnt, wird als Hauptgrund für die Abweichungen zwischen Mikrozensus und Bildungsstatistik in Bezug auf Schüler der gymnasialen Oberstufe angenommen, dass diese auf Begriffe und Definitionen zurückzuführen sind, die sich in der Befragungspraxis nicht unmittelbar umsetzen lassen. Zur Überprüfung dieser Vermutung stehen daher Vergleiche des Mikrozensus 1996 mit der Bildungsstatistik im Zentrum dieses Abschnitts. Sie werden durch Vergleiche früherer Mikrozensusdaten ergänzt.

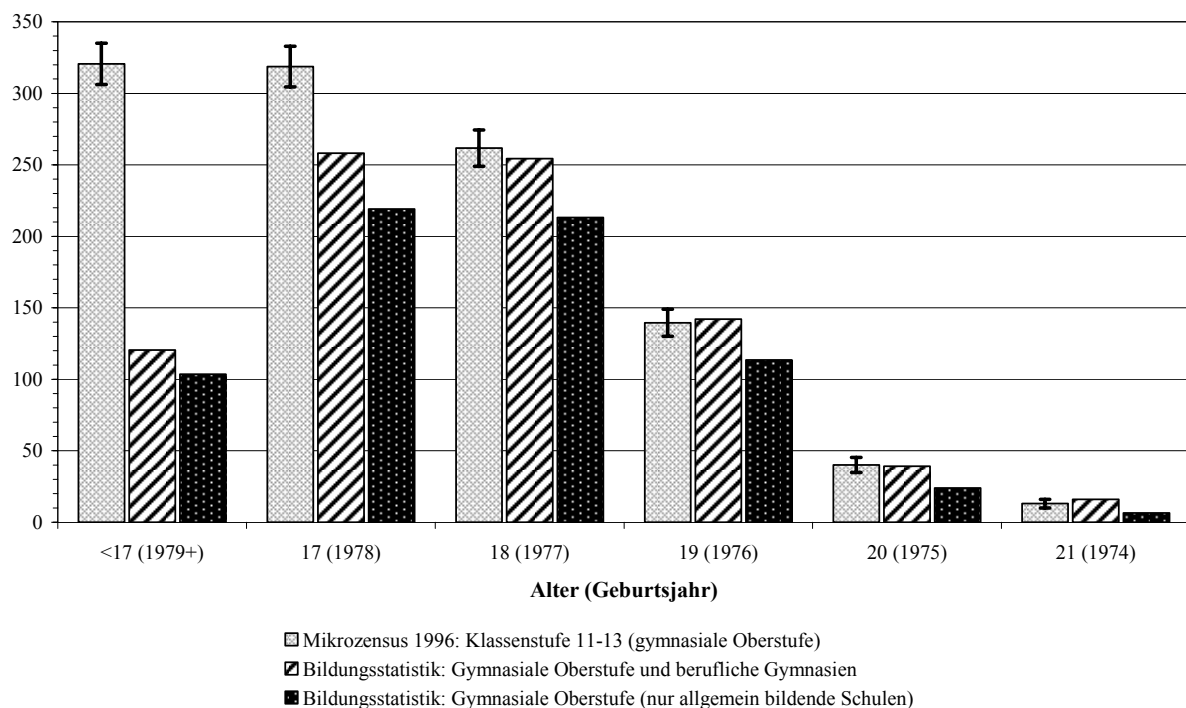
Gemäß der Bildungsstatistik besuchten im Schuljahr 1995/96 679.900 Schüler im Alter bis zu 21 Jahren die gymnasiale Oberstufe (siehe Abb. 1). Der Mikrozensus 1996 weist dagegen hochgerechnet 1.093.600 Schüler aus. Die Differenz von über 413.000 Schülern entspricht einer Übererfassung von insgesamt rund 60 Prozent. Bei den Gesamtwerten des Mikrozensus sind die 95 %-Konfidenzintervalle ausgewiesen.⁴ Die Differenzen zur Bildungsstatistik sind offensichtlich nicht durch den Stichprobenfehler zu erklären. Die Überschätzung reduziert sich insgesamt auf ungefähr die Hälfte (32 %), wenn zu den Schülern der Sekundarstufe II allgemein bildender Schulen der Bildungsstatistik die Schüler beruflicher Gymnasien hinzugerechnet werden. In diesem Fall schließt das Konfidenzintervall des Mikrozensus bei den 18- bis 21-Jährigen die entsprechenden Gesamtwerte der Bildungsstatistik ein. Dies deutet darauf hin, dass im Mikrozensus Besucher beruflicher Gymnasien entgegen den Definitionen des Mikrozensus als Schüler der gymnasialen Oberstufe erfasst sind. Für diese Annahme spricht auch, dass durch die Herausnahme der Schüler beruflicher Gymnasien (Fachoberschulen, Fachgymnasien und Berufs-/Technische Oberschulen) aus den Gesamtwerten der Berufsschulstatistik die Anpassung des Mikrozensus an die Berufsschulstatistik verbessert werden kann (hier ohne Nachweis; siehe Schimpl-Neimanns 2005: 28-29).

Bei den unter 18-Jährigen, insbesondere bei den unter 17-Jährigen bestehen dennoch weiterhin erhebliche Übererfassungen, die offensichtlich nicht mit der Fehlklassifikation beruflicher

⁴ Zur Berechnung der Stichprobenfehler bei der hier verwendeten gebundenen Hochrechnung (Anpassung an die Bevölkerungsfortschreibung) siehe Rendtel und Schimpl-Neimanns (2001: 103f.). Dieses Verfahren ist nur mit den Scientific Use Files des Mikrozensus ab 1996 anwendbar.

Gymnasien zu erklären sind. Diese Abweichungen könnten auf das oben genannte Problem des Berichtszeitraums bzw. auf die Fokussierung auf den Interviewzeitpunkt und Antworten zum Schulbesuchsstatus des neuen Schuljahres zurückzuführen sein. Sie könnten aber auch damit zusammenhängen, dass Befragten die Unterscheidung des Schulbesuchs nach Sekundarstufen (Klassenstufen 5-10 vs. 11-13) schwer fällt und die zur Klassenstufe 11-13 in Klammern genannte Ergänzung "gymnasiale Oberstufe" allgemein als "Gymnasium" missverstanden wird und von der Unterscheidung der Sekundarstufen ablenkt.⁵ Da die Bildungsstatistik die Altersgruppen, bei denen Überschneidungen zwischen den Sekundarstufen I und II auftreten, nur zusammengefasst in nach oben bzw. unten offenen Flügelklassen ausweist, lassen sich diese Vermutungen nicht weiter verfolgen.

Abbildung 1: Besucher der gymnasialen Oberstufe im Mikrozensus 1996 und in der Bildungsstatistik (Schuljahr 1995/96) in 1.000



Quellen:

Bildungsstatistik: Gymnasiale Oberstufe allgemein bildender Schulen: Statistisches Bundesamt 1997a: 24, 36-37. Berufliche Schulen: Statistisches Bundesamt 1997b: 17. (Gymnasiale Oberstufe: N=679.900; gymnasiale Oberstufe incl. berufliche Gymnasien: N=830.300).

Mikrozensus 1996 (faktisch anonymisierte 70 %-Substichprobe), Bevölkerung am Ort der Hauptwohnung; hochgerechnete, an die Bevölkerungsfortschreibung angepasste Fallzahlen (N=1.093.600; ungewichtet: n=6.753); 95 %-Konfidenzintervall: |—|; eigene Berechnungen.

⁵ Im Mikrozensus 1996 enthält die Flügelklasse der unter 17-Jährigen auch den Geburtsjahrgang 1981. Dass hochgerechnet rund 18.400 Schüler, die im Herbst 1995 etwa 14 bis 15 Jahre alt waren, einen Besuch der gymnasialen Oberstufe angeben, könnte als Indiz für eine unscharfe Differenzierung zwischen den Sekundarstufen I und II gewertet werden.

Die genannten Befunde zeigen sich auch für die Erhebungszeitpunkte 1997 bis 1999 (siehe Schimpl-Neimanns 2005: 56-58) und weisen somit auf ein systematisches Problem hin. Im dritten Abschnitt wurden als Ursachen auch andere potenzielle Zusammenhänge zwischen Befragungsart und Verteilungsabweichung angesprochen. Um diese versuchsweise zu kontrollieren, wurde eine logistische Regression zum Verhältnis des Besuchs der gymnasialen Oberstufe vs. einer beruflichen Schule mit soziodemografischen Variablen und den Befragungsarten Selbstausfüller und Proxy-Interview geschätzt. Mit den Daten des Mikrozensus 1999 ergaben sich keine statistisch signifikanten Interaktionen zwischen den Variablen der Befragungsart und den anderen erklärenden Variablen (hier ohne Darstellung). Damit kann zwar nicht ausgeschlossen werden, dass solche Effekte dennoch vorhanden sind. Für eine Überprüfung von Mode-Effekten liegen jedoch im Mikrozensus keine geeigneten Informationen vor. Von entscheidender Bedeutung wäre die Kenntnis des Interviewzeitpunkts.

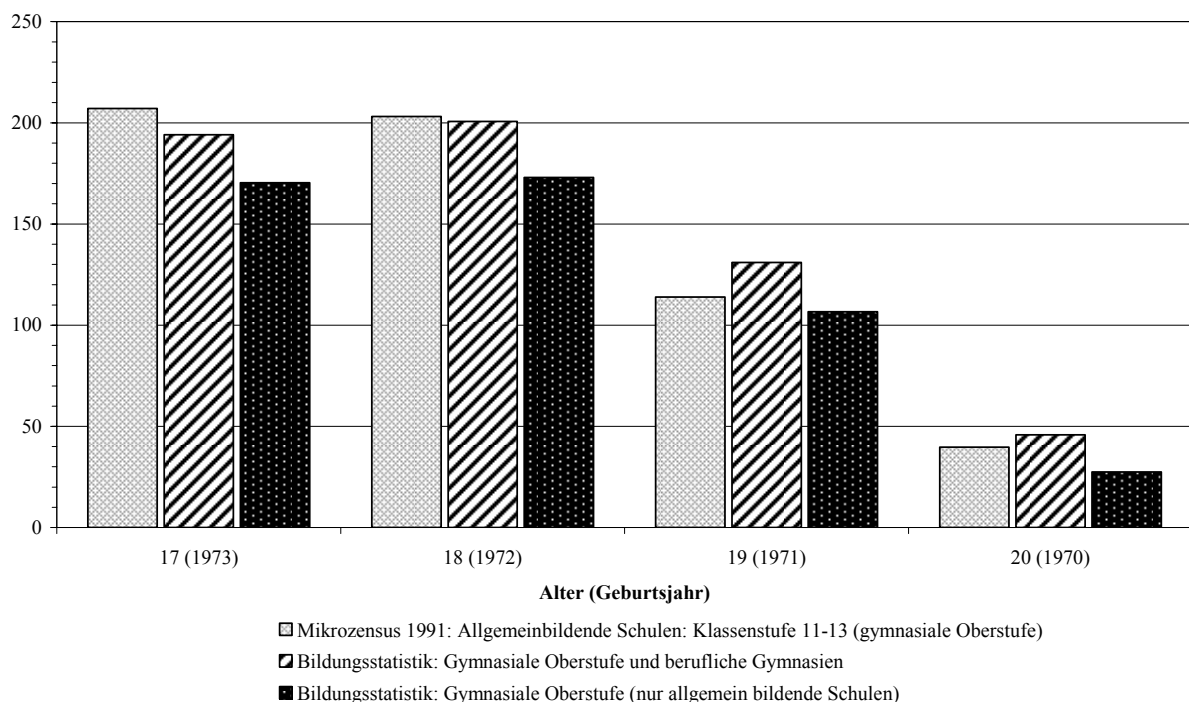
Um die Plausibilität der Erklärung der Verteilungsunterschiede trotz fehlender methodischer Informationen zu testen, kann auf den Mikrozensus 1989 zurückgegriffen werden. Bis 1990 wurde der Schulbesuch nach Schularten erfragt. Alle allgemein bildenden und beruflichen Schulen mit dem Bildungsziel allgemeine Hochschulreife (Abitur), fachgebundene Hochschulreife oder Fachhochschulreife wurden in der Kategorie "Gymnasium/Fachoberschule" zusammengefasst. Trifft die Annahme der Fehlklassifikation der beruflichen Gymnasien zu, ist zu erwarten, dass die Abweichungen zwischen Bildungsstatistik und Mikrozensus 1989 geringer ausfallen. Zum Vergleich werden zunächst die Mikrozensusergebnisse des Jahres 1991 herangezogen. Da im Mikrozensus 1989 die Abgrenzung der Sekundarstufe II nur näherungsweise mittels des Alters möglich ist, wird im Folgenden auf die nach unten offene Flügelklasse sowie auf die Altersgruppe der 21-Jährigen, die für einige Schulformen in der Bildungsstatistik ebenfalls zusammengefasste Jahrgänge enthält, verzichtet. Außerdem beziehen sich die Gegenüberstellungen auf Westdeutschland, d. h. nur auf das frühere Bundesgebiet.

Der Vergleich des Mikrozensus 1991 mit der Bildungsstatistik (siehe Abb. 2) zeigt in der Tendenz geringere, aber im Muster ähnliche Abweichungen wie beim Mikrozensus 1996.⁶ Insgesamt wird die Zahl der gymnasialen Oberstufenschüler um 18 Prozent (= 563.800 / 477.500) überschätzt. Zur Gesamtabweichung trägt hauptsächlich die Übererfassung von rund 22 Prozent bei den 17-Jährigen bei. Dies verweist wiederum auf die Probleme der Unterscheidung der Sekundarstufen bzw. die Fokussierung auf den Interviewzeitpunkt.

⁶ Da die neuen Fragen zum Schulbesuch erstmals 1991 gestellt wurden, ist nicht auszuschließen, dass die Interviewer teilweise noch die alten Definitionen zum Schulbesuch angewendet haben. Dies könnte ein Grund für die 1991 im Vergleich zu 1996 geringeren Abweichungen zwischen Mikrozensus und Bildungsstatistik und das im Vergleich zum Mikrozensus 1989 ähnliche Abweichungsmuster sein.

Wie beim Mikrozensus 1996 sind, mit Ausnahme der 19-Jährigen, die altersspezifischen Abweichungen zur Bildungsstatistik geringer, wenn zu den Besuchern der gymnasialen Oberstufe auch die Schüler beruflicher Gymnasien gezählt werden. Bei den über 18-Jährigen kehrt sich allerdings die Übererfassung in eine Untererfassung von rund 13 Prozent um. Insgesamt ist die Differenz zwischen Mikrozensus und Bildungsstatistik von 1,3 Prozent vernachlässigbar, wenn man in der Bildungsstatistik die Schüler beruflicher Gymnasien hinzunimmt.

Abbildung 2: Besucher der gymnasialen Oberstufe im Mikrozensus 1991 und in der Bildungsstatistik (Schuljahr 1990/91) in 1.000 - Westdeutschland



Quellen:

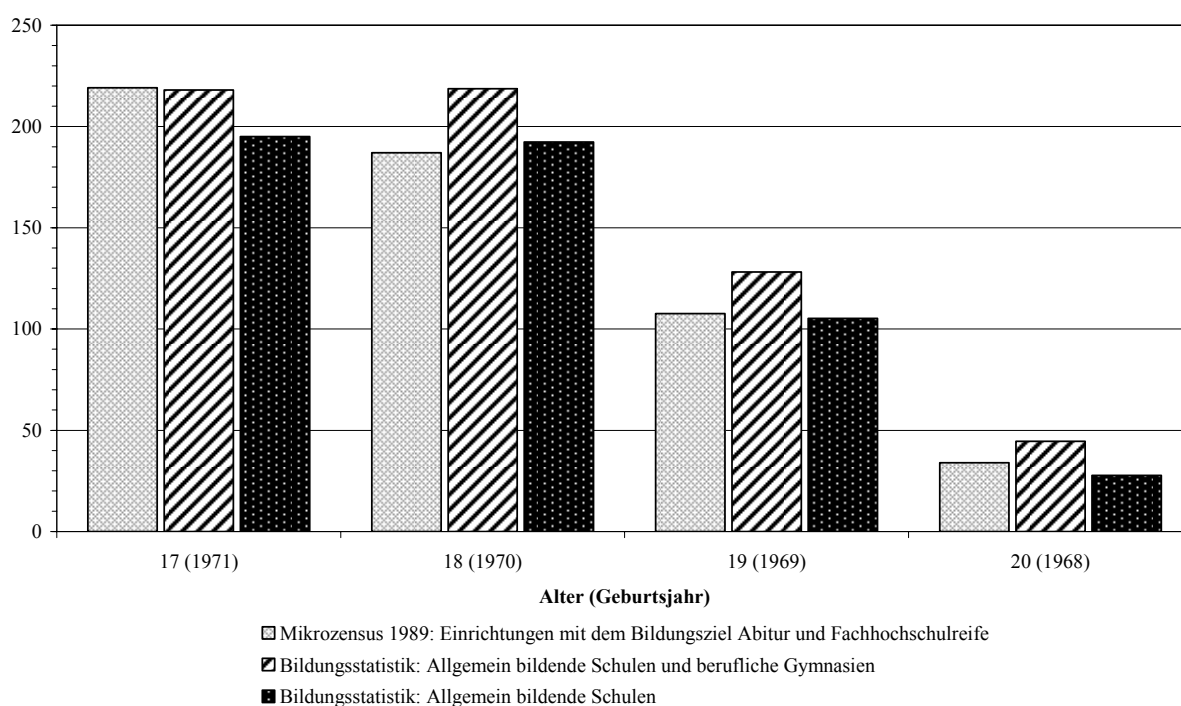
Bildungsstatistik: Gymnasiale Oberstufe allgemein bildender Schulen (N=477.500); gymnasiale Oberstufe incl. berufliche Gymnasien (N=571.500); Statistisches Bundesamt 1992.

Mikrozensus 1991 (faktisch anonymisierte 70 %-Substichprobe), Bevölkerung am Ort der Hauptwohnung; hochgerechnete, an die Bevölkerungsfortschreibung angepasste Fallzahlen (N=563.800; ungewichtet n=3.578); eigene Berechnungen.

Mit dem Mikrozensus 1989 werden 547.700 Schüler im Alter von 17 bis 20 Jahren geschätzt, die Schulen mit dem Bildungsziel des Abiturs oder der Fachhochschulreife besuchen (siehe Abb. 3). Der hierzu vergleichbare Gesamtwert der Bildungsstatistik beträgt 609.400 Schüler. Er umfasst die der gymnasialen Oberstufe zugeordneten allgemein bildenden Schulen (siehe Abschnitt 2.3) sowie berufliche Gymnasien. Damit liegt im Mikrozensus 1989 insgesamt eine Untererfassung von zehn Prozent vor. Im Vergleich zur Übererfassung der gymnasialen Oberstufe im Mikrozensus 1991 (18 %) und insbesondere zur gravierenden Übererfassung dieser Gruppe im Mikrozensus 1996 (60 %) ist die (absolute) Abweichung des Mikrozensus 1989

vom Sollwert der Bildungsstatistik deutlich geringer. Dies unterstützt die obige Vermutung, dass die im Mikrozensus ab 1991 verwendete Abgrenzung der gymnasialen Oberstufe mit der Zuordnung beruflicher Gymnasien zu beruflichen Schulen in der Befragungspraxis schwer umzusetzen ist und eine wichtige Quelle der Verteilungsabweichungen zur Bildungsstatistik darstellt.

Abbildung 3: Schüler allgemein bildender Schulen und beruflicher Gymnasien mit dem Bildungsziel Abitur und Fachhochschulreife im Mikrozensus 1989 und in der Bildungsstatistik (Schuljahr 1988/89) in 1.000



Quellen:

Bildungsstatistik: Allgemein bildende Schulen (N=520.400); allgemein bildende Schulen incl. berufliche Gymnasien (N=609.400); Statistisches Bundesamt 1990.

Mikrozensus 1989 (faktisch anonymisierte 70 %-Substichprobe), Bevölkerung am Ort der Hauptwohnung; hochgerechnete, an die Bevölkerungsfortschreibung angepasste Fallzahlen (N=547.700; ungewichtet n=3.338); eigene Berechnungen.

In Bezug auf die Frage, ob die Abweichungen zwischen dem Mikrozensus ab 1991 und der Bildungsstatistik bei den unter 18-Jährigen mit dem Problem der Fokussierung auf den Interviewzeitpunkt oder mit der schwierigen Unterscheidung der Sekundarstufen I und II (bzw. Klassenstufen 5-10 vs. 11-13) zusammenhängen, ist der Vergleich der Verteilungsunterschiede der Mikrozensusen 1989 und 1991 zur Bildungsstatistik aufschlussreich. Verwendet man als Sollzahlen der Bildungsstatistik die Gesamtwerte der Schüler allgemein bildender Schulen der gymnasialen Oberstufe einschließlich beruflicher Gymnasien, zeigt sich für die 17-Jährigen folgender Befund: Während im Mikrozensus 1991 eine Übererfassung des Ge-

samtwerts der Bildungsstatistik um rund sieben Prozent vorliegt,⁷ wird der Sollwert der Bildungsstatistik mit dem Mikrozensus 1989 nur geringfügig um 0,5 Prozent überschätzt; dies ist vernachlässigbar. Da in beiden Mikrozensus-erhebungen nach dem „gegenwärtigen Schulbesuch“ gefragt wird, wäre davon auszugehen, dass Effekte der Fokussierung auf den Interviewzeitpunkt gleichermaßen sowohl 1989 als auch 1991 auftreten. Wenn aber bei den 17-Jährigen 1989 im Unterschied zu 1991 keine Übererfassung festzustellen ist, kann dies 1989 nicht mit Fokussierungseffekten zusammenhängen. Vergleichbar zum Mikrozensus 1989 enthalten die Zahlen der Bildungsstatistik des Schuljahres 1988/89 neben den 17- bis 20-jährigen Schülern beruflicher Gymnasien die gleichaltrigen Schüler allgemein bildender Schulen, die der gymnasialen Oberstufe zuzurechnen sind. Zwar wird dabei die Sekundarstufe II nur näherungsweise durch die Altersgliederung abgegrenzt, dennoch spricht das Ergebnis dafür, dass die in den Mikrozensus ab 1991 bei den jüngeren, unter 18-jährigen Schülern der gymnasialen Oberstufe festgestellten starken Übererfassungen kaum auf Fokussierungseffekte, sondern auf eine nicht adäquate Unterscheidung der Sekundarstufen I und II zurückzuführen sind.

Die Untererfassung der 17- bis 20-jährigen Schüler allgemein bildender Schulen der gymnasialen Oberstufe und beruflicher Gymnasien im Mikrozensus 1989 um zehn Prozent ist jedoch nicht zu vernachlässigen. Bei der Frage, womit diese Untererfassung zusammenhängen könnte, ist an die Anmerkungen in Abschnitt 2.3 anzuknüpfen. Wie dort ausgeführt, werden die genauen Zuordnungen der Schulformen zu den Antwortkategorien im Mikrozensus 1989 nur im Interviewerhandbuch genannt. Sowohl im Interviewerbogen als auch im Selbstaussfüllerbogen ist lediglich der Sammelbegriff „Gymnasium/Fachoberschule“ angegeben. Nicht explizit genannte Schulen können auf diese Weise leicht übersehen werden. Für die Untererfassung dürfte somit maßgeblich die unvollständige Antwortliste verantwortlich sein.

5. Antwortstabilität der Angaben zum allgemeinen Schulabschluss

Wie wir bisher gesehen haben, treten auch im Mikrozensus systematische Fehler in Form von Klassifikationsfehlern etc. auf. Die Panelangaben des Mikrozensus eröffnen neue Möglichkeiten zur Beurteilung der Antwortkonsistenz bzw. Antwortstabilität im Zeitverlauf. Dabei können geringe Antwortstabilitäten ergänzend zu den vorherigen Analysen auf Probleme bei der Erhebung von Bildungsabschlüssen hinweisen. Trifft die zum Schulbesuch gezogene Schlussfolgerung zu, dass Schüler der gymnasialen Oberstufe und beruflicher Schulen nicht

⁷ Beim Mikrozensus 1996 (einschließlich neue Bundesländer) beträgt die Übererfassung der 17-jährigen Schüler der gymnasialen Oberstufe einschließlich beruflicher Gymnasien rund 23 Prozent.

den Definitionen des Mikrozensus entsprechend unterscheidbar sind, ist zu erwarten, dass sich dies auch bei den in diesen Schulformen erworbenen typischen Abschlüssen der allgemeinen bzw. fachgebundenen Hochschulreife (gymnasiale Oberstufe) vs. Fachhochschulreife (berufliche Schulen) widerspiegelt. Aus Mustern von Übergangswahrscheinlichkeiten lässt sich allgemein schließen, dass den Befragten bestimmte Antwortkategorien ähnlich erscheinen. Demnach sind beim Fachhochschulreifeabschluss im Vergleich zu anderen Abschlüssen eine geringere zeitliche Stabilität und eine höhere Übergangswahrscheinlichkeit zum Abitur zu erwarten.

Zur Qualität von Bildungsangaben kann nur auf sehr wenige Erkenntnisse zurückgegriffen werden (Reimer 2001: 114). In der Test-Retest Studie zum ALLBUS 1984 (n=154) wurden drei Interviews im Abstand von vier Wochen durchgeführt. Für den allgemeinen Schulabschluss wurde eine Stabilität von 89 Prozent, für die berufliche Qualifikation nur eine Stabilität von 72 Prozent ermittelt (Porst und Zeifang 1987: 196). Auf der Basis einer zweimaligen Befragung 89 verheirateter und geschiedener Ehepaare im Abstand von etwa einem halben Jahr, mit der sowohl Proxyeffekte als auch Rückerinnerungseffekte untersucht wurden, berichtet Babka von Gostomki (1995) für den Schulabschluss eine Gesamtstabilität von 72 Prozent. Orientiert man sich an diesen zwei Studien, kann eine Antwortstabilität von 70 Prozent als Minimum bezeichnet werden.

Mit dem Mikrozensuspanel liegt für diese Fragestellungen erstmals ein wesentlich umfangreicherer Datensatz vor (vgl. Herter-Eschweiler 2003: 243f.). Einschränkend muss jedoch bemerkt werden, dass keine Information über die Auskunft gebende Person verfügbar ist. Die Proxy-Angaben wurden zwar erstmals 1999 im Rahmen der EU-Arbeitskräftestichprobe erhoben, jedoch aus Datenschutzgründen nicht in das File aufgenommen. Im Folgenden werden nur räumlich immobile Personen im Alter von 25 bis 50 Jahren ohne Schulbesuch betrachtet, bei denen man von einer abgeschlossenen Schulkarriere ausgehen kann. Mit der oberen Altersgrenze werden der Einfachheit halber die besonderen Aspekte der Freistellung von der Auskunftspflicht für über 51-Jährige ausgeblendet. Ebenfalls vereinfacht die Konzentration auf räumlich Immobile bzw. der Ausschluss wegziehender Personen die Ergebnisdarstellung. Die Selektion dürfte für die Frage konsistenter Antworten unerheblich sein. Für rund 29.000 Personen werden insgesamt rund 87.000 Übergänge beobachtet, d. h. aufgrund von zeitlich inkonsistenten Angaben liegen Mehrfachantworten vor.

Tabelle 1 zeigt für den Abschluss der allgemeinen bzw. fachgebundenen Hochschulreife, dem typischen Abschluss der gymnasialen Oberstufe, eine Antwortstabilität von 85 Prozent. Für

Personen mit Fachhochschulreife, die zumeist an beruflichen Schulen erworben wird, wird lediglich eine Wahrscheinlichkeit von 50 Prozent festgestellt, dass dieser Abschluss auch bei der nächsten Befragung genannt wird. Die Übergangswahrscheinlichkeit zur Angabe Abitur beträgt 23 Prozent. Diese überdurchschnittlich inkonsistenten Abschlussangaben korrespondieren mit der oben bereits festgestellten Schwierigkeit von Befragten, den Besuch beruflicher und allgemein bildender Schulen der Sekundarstufe II zu differenzieren.

Tabelle 1: Antwortvariabilitäten zum allgemeinen Schulabschluss im Mikrozensuspanel 1996-1999 (Übergangswahrscheinlichkeiten zwischen den Zeitpunkten t und t+1) – Zeilenprozentwerte

Bildungsangabe (t) (m Beobachtungen; in Prozent)	Bildungsangabe (t+1)							
	0	1	2	3	4	5	8	9
0 Kein Abschluss (1,7 %)	46,8	40,0	2,0	2,8	0,3	1,6	1,1	5,5
1 Volks-/Hauptschule (40,3 %)	1,5	87,4	1,8	5,3	0,5	0,7	0,5	2,4
2 Polytechnische Oberschule (14,6 %)	0,2	5,9	83,9	5,1	1,5	1,5	0,2	1,7
3 Realschule (20,1 %)	0,2	10,0	3,1	79,0	3,0	2,4	0,4	1,9
4 Fachhochschulreife (4,2 %)	0,1	4,7	5,3	13,9	50,5	23,3	0,3	1,9
5 Allgem./fachgeb. Hochschulreife (15,2 %)	0,2	1,6	1,3	3,2	6,5	85,2	0,4	1,7
8 Abschluss vorh., ohne Angabe (0,5 %)	4,4	41,3	5,4	17,6	3,9	9,1	7,2	11,1
9 Keine Angabe (3,3 %)	2,8	36,6	8,3	14,4	3,0	9,8	2,6	22,4
Insgesamt (m=87.234; 100 %)	1,6	40,9	14,4	20,4	4,3	15,1	0,5	2,8

Quelle: Faktisch anonymisiertes Mikrozensuspanel 1996-1999; räumlich immobile Personen im Alter von 25 bis 50 Jahren ohne Schulbesuch (n=29.078 Personen).

Hohe Antwortkonsistenzen können des Weiteren für die Abschlüsse der Volks- bzw. Hauptschule, der Polytechnischen Oberschule und der Realschule festgestellt werden. Personen, die zunächst angegeben haben, keinen allgemein bildenden Abschluss zu besitzen bzw. ihren Abschluss nicht genannt haben, wechseln im Erhebungszeitraum 1996-1999 häufig zur Angabe eines Hauptschulabschlusses (siehe Kategorien 0, 8 und 9 in Tab. 1). Auch wenn dies wegen der zugrunde liegenden kleinen Fallzahlen nur eingeschränkt zu interpretieren ist, deutet es darauf hin, dass die Panelangaben zur Korrektur des Item-Nonresponse bzw. zur Überprüfung der Angabe "kein Abschluss" genutzt werden können.

6. Zusammenfassung

Die am Beispiel des Besuchs der gymnasialen Oberstufe untersuchten Fragen zur Datenqualität haben auf einige Problemstellen aufmerksam gemacht. Das wohl überraschendste Ergebnis besteht in der erheblichen Übererfassung des Besuchs der gymnasialen Oberstufe (Klas-

senstufe 11-13) durch den Mikrozensus 1996 um insgesamt rund 60 Prozent im Vergleich zur Bildungsstatistik. Diese Übererfassung der Oberstufenschüler ist auf Klassifikationsfehler zurückzuführen, die mit schwierigen, dem Alltagsverständnis nicht direkt entsprechenden Definitionen im Mikrozensus zusammenhängen. Konkret betrifft es die Unterscheidung zwischen den Klassenstufen 11-13 allgemein bildender Schulen („gymnasiale Oberstufe“) einerseits und beruflichen Schulen andererseits, zu denen laut Mikrozensus auch berufliche Gymnasien zählen. Offensichtlich werden die Definitionen und Antwortvorgaben des Mikrozensus von den Befragten nicht vollständig verstanden. Die Gegenüberstellungen mit der Bildungsstatistik deuten darauf hin, dass im Mikrozensus Schüler beruflicher Gymnasien als gymnasiale Oberstufenschüler erfasst sind. Die gravierende Übererfassung der unter 18-jährigen Oberstufenschüler hängt vermutlich damit zusammen, dass die Unterscheidung der Klassenstufen 5-10 vs. 11-13 bzw. der Sekundarstufen I und II nicht klar ist. Zwar sind diese Befunde teilweise spekulativ, aber auch plausibel. Für alternative Erklärungen und weitere Kontrollen wären zusätzliche Informationen nötig.

Korrespondierend zu der in der Befragungspraxis offenbar schwierigen Unterscheidung zwischen gymnasialer Oberstufe und beruflichen Gymnasien zeigen sich diese Probleme auch in Bezug auf die Konsistenz bzw. Stabilität der Antworten zum allgemeinen Schulabschluss. Für das i. d. R. mit Abschluss der gymnasialen Oberstufe erworbene Abitur kann mit 85 Prozent eine sehr hohe Stabilität festgestellt werden. Dagegen liegt die Stabilität des beruflichen Abschlusses Fachhochschulreife lediglich bei 50 Prozent. Für die anderen Kategorien belegen die ersten hier vorgestellten Ergebnisse des Mikrozensuspanels zur Konsistenz der Antworten zum allgemeinen Schulabschluss erfreulicherweise hohe Antwortstabilitäten von rund 80 Prozent und mehr.

So ernüchternd die Einbußen der Qualität der Angaben zum Schulbesuch für Bildungsforscher auch sein mögen, ist es gleichwohl für jede sachgerechte Auswertung wichtig zu wissen, in welcher Weise die Daten systematische Fehler aufweisen. Es geht aber nicht nur um die Übereinstimmung mit externen, für valider gehaltenen Datenquellen per se. Für einfache Populationsschätzungen werden die Mikrozensusdaten kaum benötigt, hierfür kann die Bildungsstatistik als Totalerhebung direkt verwendet werden. Die Vorteile des Mikrozensus bestehen vielmehr in der Verfügbarkeit weiterer Merkmale für differenzierte deskriptive Analysen und statistische Modelle. Bei der Verwendung von Bildungsmerkmalen in statistischen Modellen wird infolge fehlender Informationen zur Datenqualität zumeist angenommen, dass die Merkmale frei von systematischen Fehlern sind. In dieser Hinsicht können auf Grundlage der hier berichteten Befunde erste und einfache Konsequenzen gezogen werden.

Beispielsweise kann nun bei Analysen zum Besuch der gymnasialen Oberstufe darauf verwiesen werden, dass diese Kategorie nicht nur allgemein bildende Schulen umfasst. Durch die Beschränkung auf über 17-Jährige lässt sich außerdem eine gute Anpassung der Mikrozensusergebnisse an die Bildungsstatistik erreichen, wenn zur gymnasialen Oberstufe auch berufliche Gymnasien gezählt werden.

In diesem Zusammenhang der Verwendung bildungsstatistischer Merkmale des Mikrozensus in der Forschung ist darauf hinzuweisen, dass die aktuellen Fragen zum Schulbesuch ab dem Mikrozensus 2005 geändert wurden. Erfragt wird nun der Schulbesuch in den letzten vier Wochen sowie in den letzten zwölf Monaten vor der jeweiligen Befragung. Zwar wird im Fragebogen nach wie vor zwischen dem Besuch allgemein bildender vs. beruflicher Schulen unterschieden, jedoch wird ab 2005 zum Besuch einer beruflichen Schule explizit auch der jeweils erreichbare Abschluss genannt. Ob damit schon die hier festgestellten Fehlklassifikationen der Besucher beruflicher Gymnasien reduziert werden können, bleibt weiteren Analysen vorbehalten. Vor dem Hintergrund, dass der Schulbesuch in den einzelnen Bundesländern unterschiedlich geregelt ist, wird die Erhebung des Schulbesuchs und von Bildungsabschlüssen nicht nur im Mikrozensus, sondern vermutlich auch in anderen bundesweiten Umfragen auch weiterhin eine Herausforderung darstellen.

Abschließend ist festzuhalten, dass in Deutschland insbesondere im Vergleich zu den in den USA routinemäßig durchgeführten Analysen allgemein zur Datenqualität und besonders zu Messfehlern ein erheblicher Rückstand besteht. Die europäischen statistischen Ämter unternehmen seit einiger Zeit verstärkt Anstrengungen um diese Defizite aufzuholen (Blanc et al. 2001). Als ein Ergebnis dieser Orientierung auf Aspekte der Datenqualität stellt das Statistische Bundesamt seit Kurzem zu verschiedenen Erhebungen und Statistiken Qualitätsberichte zur Verfügung (Körner und Schmidt 2006).⁸ Aus Sicht der Forschung ist es sehr wünschenswert, dass diesem Schritt weitere, differenziertere Untersuchungen folgen. Die Gründe dafür sind in Empfehlungen an die Datenproduzenten zusammengefasst (Office of Management and Budget 2001: 6-25): „(...) a data user cannot understand the limitations of the data – from a measurement error point of view – unless the data collection program takes steps to explicitly provide such information.“ Als Nutzer amtlicher Mikrodaten ist aber auch die akademische Forschung gefordert. In dieser Hinsicht wird das Scientific Use File des Mikrozensuspanels – neben neuen substanzwissenschaftlich orientierten Verlaufsanalysen – für Fragen zur Datenqualität eine Vielzahl weiterer methodischer Auswertungsmöglichkeiten bieten.

⁸ Siehe unter www.destatis.de/allg/d/veroe/qualitaetsberichte.htm.

Literatur

- Afentakis, Anja, und Wolf Bihler*, 2005: Das Hochrechnungsverfahren beim unterjährigen Mikrozensus ab 2005. *Wirtschaft und Statistik* (10): 1039-1048.
- Babka von Gostomski, Christian*, 1995: Zur Konsistenz und Übereinstimmung von Ehepartnern bei retrospektiv erhobenen Angaben zur Person und zur Beziehung. *Zeitschrift für Familienforschung* 7(1): 6-26.
- Bellenberg, Gabriele, Gertrud Hovestadt und Klaus Klemm*, 2004: Selektivität und Durchlässigkeit im allgemein bildenden Schulsystem. Universität Duisburg-Essen: Arbeitsgruppe Bildungsforschung/Bildungsplanung.
- Biemer, Paul M., und Lars E. Lyberg*, 2003: *Introduction to Survey Quality*. Hoboken, New Jersey: Wiley.
- Black, Dan, Seth Sanders und Lowell Taylor*, 2003: Measurement of Higher Education in the Census and CPS. *Journal of the American Statistical Association* 98(463): 545-554.
- Blanc, Michel, Walter Radermacher und Thomas König*, 2001: Qualität und Nutzer: Grundlagen und Instrumente der Nutzerorientierung in der amtlichen Statistik. *Wirtschaft und Statistik* (10): 799-807.
- Breiholz, Holger*, 2000: Ergebnisse des Mikrozensus 1999. *Wirtschaft und Statistik* (5): 328-336.
- Brick, J. Michael, Jacque Wernimont und Maida Montes*, 1996: The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component. NCES Working Paper No. 96-14. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Dawe, Fiona, und Ian Knight*, 1997: A study of proxy response on the Labour Force Survey. *Survey Methodology Bulletin* 40: 30-36.
- Dräther, Hendrik, Uwe Fachinger und Angelika Oelschläger*, 2001: Selbständige und ihre Altersvorsorge – Möglichkeiten der Analyse anhand der Mikrozensus und erste Ergebnisse. ZeS-Arbeitspapier Nr. 1/01. Universität Bremen: Zentrum für Sozialpolitik.
- Esser, Hartmut, Heinz Grohmann, Walter Müller und Karl-August Schäffer*, 1989: Mikrozensus im Wandel. Untersuchungen und Empfehlungen zur inhaltlichen und methodischen Gestaltung. In: Statistisches Bundesamt (Hg.): *Forum der Bundesstatistik*, Band 11. Stuttgart: Metzler-Poeschel.
- Heidenreich, Hans-Joachim*, 1994: Hochrechnung des Mikrozensus ab 1990. S. 112-123 in: *Siegfried Gabler, Jürgen Hoffmeyer-Zlotnik und Dagmar Krebs* (Hg.): *Gewichtung in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Herter-Eschweiler, Robert*, 2003: Längsschnittdaten aus dem Mikrozensus. Basis für neue Analysemöglichkeiten. Bonn: Statistisches Bundesamt (unveröffentlichtes Manuskript).
- KMK [Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland], 2001: Grundstruktur des Bildungswesens in der Bundesrepublik Deutschland. Diagramm 1999 für Faltblatt, Sonderdruck und Download. Bonn. URL: www.kmk.org/schul/home.htm - .../doku/ddiagr.doc.
- KMK, 2003: Das Bildungswesen in der Bundesrepublik Deutschland 2002. Bonn. URL: www.kmk.org/dossier/dossier_2002/dossier_2002_dt_ebook.pdf.
- Körner, Thomas, und Jürgen Schmidt*, 2006: Qualitätsberichte – ein neues Informationsangebot über Methoden, Definitionen und Datenqualität der Bundesstatistiken. *Wirtschaft und Statistik* (2): 109-117.

- Lois, Daniel*, 2005: Weiterbildungsbeteiligung älterer Erwerbstätiger – Die Messung im Mikrozensus und der Einfluss soziodemografischer Variablen. Beitrag zur 4. Nutzerkonferenz „Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und zum Arbeitsmarkt“, Mannheim, 12./13.10.2005. URL: www.gesis.org/Dauerbeobachtung/GML/Service/Veranstaltungen/4.NK_2005/papers/18_Lois.pdf.
- Lotze, Sabine*, und *Holger Breiholz*, 2002a: Zum neuen Erhebungsdesign des Mikrozensus. Teil 1. *Wirtschaft und Statistik* (5): 359-366.
- Lotze, Sabine*, und *Holger Breiholz*, 2002b: Zum neuen Erhebungsdesign des Mikrozensus. Teil 2. *Wirtschaft und Statistik* (6): 454-459.
- McGovern, Pamela D.*, und *John M. Bushery*, 1999: Data mining the CPS reinterview: Digging into response error. URL: www.fcs.gov/99papers/mcgovern.pdf.
- Meyer, Kurt*, 1994: Zum Auswahlplan des Mikrozensus ab 1990. S. 106-111 in: *Siegfried Gabler, Jürgen Hoffmeyer-Zlotnik und Dagmar Krebs* (Hg.): *Gewichtung in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- NCES* [U.S. Department of Education, National Center for Education Statistics], 1997: *Measurement Error Studies at the National Center for Education Statistics*, NCES 97-464. Washington, DC: NCES.
- Office of Management and Budget*, 2001: *Measuring and Reporting Sources of Error in Surveys*. Statistical Policy Working Paper 31. URL: www.fcs.gov/01papers/SPWP31_final.pdf.
- Pöschl, Hannelore*, 1992: Geringfügige Beschäftigung 1990. Ergebnisse des Mikrozensus. *Wirtschaft und Statistik* (3): 166-170.
- Porst, Rolf*, und *Klaus Zeifang*, 1987: A Description of the German General Social Survey Test-Retest Study and a Report on the Stabilities of the Sociodemographic Variables. *Sociological Methods & Research* 15(3): 177-218.
- Reimer, Maike*, 2001: Die Zuverlässigkeit des autobiographischen Gedächtnisses und die Validität retrospektiv erhobener Verlaufsdaten. Kognitive und erhebungspragmatische Aspekte. Materialien aus der Bildungsforschung Nr. 71. Berlin: Max-Planck-Institut für Bildungsforschung.
- Riede, Thomas*, und *Dieter Emmerling*, 1994: Analysen zur Freiwilligkeit der Auskunftserteilung im Mikrozensus. Sind Stichprobenergebnisse bei freiwilliger Auskunftserteilung verzerrt? *Wirtschaft und Statistik* (9): 733-742.
- Rendtel, Ulrich*, und *Bernhard Schimml-Neimanns*, 2001: Die Berechnung der Varianz von Populationsschätzern im Scientific Use File des Mikrozensus ab 1996. *ZUMA-Nachrichten* 48: 85-116.
- Rudolph, Helmut*, 1998: „Geringfügige Beschäftigung“ mit steigender Tendenz. Erhebungskonzepte, Ergebnisse und Interpretationsprobleme der verfügbaren Datenquellen. IAB-Werkstattbericht Nr. 9. Nürnberg: IAB.
- Särndal, Carl-Erik, Bengt Swensson und Jan Wretman*, 1997: *Model Assisted Survey Sampling*. New York: Springer.
- Schimml-Neimanns, Bernhard*, 2005: *Bildungsverläufe im Mikrozensuspanel 1996-1999: Besuch der gymnasialen Oberstufe bis zum Abitur*. ZUMA-Arbeitsbericht 2005/02. Mannheim: ZUMA.
- Schupp, Jürgen, Joachim Frick, Lutz Kaiser und Gert Wagner*, 1999: Zur Erhebungsproblematik geringfügiger Beschäftigung. Ein Vergleich des Mikrozensus mit dem Sozio-ökonomischen Panel und dem Europäischen Haushaltspanel. S. 93-118 in: *Paul Lüttinger*, (Hg.): *Sozialstrukturanalysen mit dem Mikrozensus*. ZUMA-Nachrichten Spezial, Band 6. Mannheim: ZUMA.

- Statistische Ämter des Bundes und der Länder*, 2004: Ergebnisse des Zensus 2001. Wirtschaft und Statistik (8): 813-833.
- Statistisches Bundesamt*, 1990: Fachserie 11 Bildung und Kultur, Reihe 1 Allgemeinbildende Schulen. Schuljahr 1988/89. Stuttgart: Metzler-Poeschel.
- Statistisches Bundesamt*, 1992: Fachserie 11 Bildung und Kultur, Reihe 1 Allgemeinbildende Schulen. Schuljahr 1990/91. Stuttgart: Metzler-Poeschel.
- Statistisches Bundesamt*, 1996: Fachserie 11 Bildung und Kultur, Reihe 1 Allgemeinbildende Schulen. Schuljahr 1995/96. Stuttgart: Metzler-Poeschel.
- Statistisches Bundesamt*, 1997a: Arbeitsunterlage Allgemeinbildende Schulen Schuljahr 1996/97. Ergänzende Tabellen zur Fachserie 11 Bildung und Kultur, Reihe 1 - Allgemeinbildende Schulen. Wiesbaden.
- Statistisches Bundesamt*, 1997b: Arbeitsunterlage Berufliche Schulen 1995/96. Ergänzende Tabellen zur Fachserie 11 Bildung und Kultur, Reihe 2 - Berufliche Schulen Schuljahr 1995/96. Wiesbaden.
- U.S. Census Bureau*, 2004: Meeting the 21st Century Demographic Data Needs - Implementing the American Community Survey. Report 9: Comparing Social Characteristics With Census 2000. Washington, CD: U.S. Census Bureau.
- Weishaupt, Horst, und Detlef Fickermann*, 2001: Informationelle Infrastruktur im Bereich Bildung und Kultur. Expertise für die Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik. In: *Kommission zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik* (Hg.): Wege zu einer besseren informationellen Infrastruktur. Baden-Baden: Nomos [CD-ROM Beilage zur Buchausgabe].