

Univ.-Prof. Dr. Gert G. Wagner

Vorsitzender der Zensuskommission

Vorsitzender des Rats für Sozial- und Wirtschaftsdaten (RatSWD)

**Materialien für das
erweiterte Berichterstattergespräch am 20. April 2009
zum
Zensusanordnungsgesetz**

Die Zensuskommission hält ihre Stellungnahme der Zensuskommission zum Merkmalskatalog und einigen anderen Details im Kabinettsentwurf des Zensusanordnungsgesetz (ZensusG2011), die sie am am 22. Januar 2009 verabschiedet hat, uneingeschränkt aufrecht, da sich die sachliche n Bedingungen, auf deren Grundlage die Stellungnahme beschlossen wurde, nicht verändert haben. Siehe

http://www.zensuskommission.de/docs/Stellungnahme_2009_01_22.pdf.

Die Zensuskommission betont in diesem Zusammenhang nochmals, dass die Beschränkung der Erhebungsmerkmale auf den EU-Pflichtkatalog (und ggf. minimale Ausweitungen) aus fachstatistischer Sicht bedauerlich ist.

Die Zensuskommission würdigt freilich ausdrücklich, dass ihre ursprünglichen Empfehlungen und Begründungen in der Begründung des Gesetzesentwurfes sorgfältig dokumentiert sind. Eine derart vorbildliche und sachgerechte Anhörung und Darstellung sollte für künftige Gesetzgebungsverfahren im Bereich der amtlichen Statistik Vorbildcharakter haben.

Das Argument der Bundesregierung, dass die Aufnahme der Merkmale zu einer erheblichen zusätzlichen Belastung der Bevölkerung und der Verwaltungen führt, ist für die Zensuskommission nicht nachvollziehbar. Die Kommission weist auch ausdrücklich darauf hin, dass der Verzicht auf diese Merkmale und Daten erhebliche Kosten verursachen wird, da zentrale Informationen, die zur Erfüllung staatlicher Aufgaben notwendig sind, nicht vorhanden sein werden.

Die Zensuskommission erkennt freilich an, dass der Gesetzgeber in der Tat frei und legitimiert ist zu entscheiden, was er als Merkmalskatalog eines Zensus für sinnvoll hält. Im Zusammenhang mit der Beschränkung des Merkmalskatalogs erinnert die Zensuskommission heute den Gesetzgeber daran, dass ein minimales Erhebungsprogramm für den Zensus ausdrücklich nur vor dem Hintergrund des von der Bundesregierung grundsätzlich zu recht vorgebrachten Arguments vertretbar ist, dass in der Bundesrepublik Deutschland der Zensus eine weniger wichtige Rolle spielt als in vielen anderen Ländern, da in der Bundesrepublik Deutschland mit dem Mikrozensus (MZ) eine laufende Erhebung permanent wesentliche Ergebnisse auf Basis einer großen Stichprobe liefert.

Da die Bundesregierung ausdrücklich die große Bedeutung des Mikrozensus (MZ) für die statistische Infrastruktur in Deutschland betont, geht die Kommission davon aus und empfiehlt dies auch heute ausdrücklich, dass die von ihr für den Zensus vorgeschlagenen Zusatzmerkmale – soweit sie noch nicht oder nicht mehr Bestandteile des Mikrozensus Erhebungsprogramms sind – unverzüglich in den Pflicht-Teil des Mikrozensus aufgenommen werden. Eine Aufnahme in den freiwilligen Teil des MZ würde den Qualitätsansprüchen, die an diese zentralen Zusatzmerkmale gestellt werden, nicht entsprechen.

Darüber hinaus betont die Zensuskommission heute ausdrücklich, dass aufgrund der überragenden Bedeutung des Mikrozensus der Gesetzgeber sicherstellen muss, dass nach Auslaufen des derzeit gültigen MZ-Gesetzes eine Weiterführung des MZ gewährleistet sein muss.

Die Zensuskommission betont darüber hinaus heute ausdrücklich, dass die bereits in der Stellungnahme vom 22. Januar thematisierte Unsicherheit (Abschnitt „Datenzugang“) bei der Interpretation des §16 des Bundesstatistikgesetzes bislang nicht ausgeräumt wurde, wodurch die Gefahr entstehen könnte, dass die Mikrodaten des Zensus nicht intensiv zur Zwecke der Qualitätssicherung künftiger Zensen re-analysiert werden könnten, weil die entsprechenden Daten der Wissenschaft nicht oder nicht auf Dauer zugänglich gemacht werden könnten. Die Zensuskommission weist auf diese Gefahr hin und bittet die Bundesregierung um entsprechende Klärung der Interpretation des §16 BStG und den Gesetzgeber ggf. um die Schaffung einer Rechtsgrundlage für die qualitätssichernde unabhängige wissenschaftliche Forschung.

Eng mit der Frage auf welche Daten qualitätssichernde Forschung zugreifen kann verbunden ist die Frage, ob die anonymisierten Mikrodaten des Zensus nur in der Abgrenzung der Merkmale, die an Eurostat geliefert werden, der Forschung zur Verfügung stehen; oder auch – was für Analysen, die sich nur auf Deutschland beziehen – in der Original-Abgrenzung der Erhebungsmerkmale.

Im folgenden wird die Bedeutung des datenschutzrechtlich einwandfreien Zugangs der Forschung zu den Mikrodaten des Zensus näher begründet. Ein ausführliches Papier, das der folgenden Stellungnahme zugrunde liegt, von Dr. Heike Wirth, Mitglied der Zensuskommission, liegt auf Nachfrage vor.

Zugang der Forschung zu den Mikrodaten des Zensus 2011

Eine wesentliche Dimension des gesellschaftlichen Nutzens der Zensusdaten ist ein an den Forschungsbedürfnissen orientierter Datenzugang, der unabhängig von der amtlichen Feststellung der Bevölkerungszahl gegeben ist. Für diesen Zweck sind sowohl der Zeitpunkt wie auch der Umfang der Datenbereitstellung wichtig. Diese Punkte werden im

Zensusanordnungsgesetz nicht geregelt und bieten damit einen erheblichen Gestaltungsspielraum, der sich für die Forschung in der Praxis als negativ erweisen kann. Die Zensuskommission artikuliert deswegen die grundlegenden Bedürfnisse der Forschung bereits zum jetzigen Zeitpunkt, mit dem Ziel, innerhalb der datenschutzrechtlichen Rahmenbedingungen einen möglichst optimalen Datenzugang und optimale Datenqualität zu erzielen. Im Folgenden werden diese Punkte näher erläutert.

Da es sich beim Zensus um eine Bundesstatistik handelt, die dem Bundesstatistikgesetz unterliegt, wird der Datenzugang demnach unter den Randbedingungen des §16(6) BStatG erfolgen. Danach dürfen *“für die Durchführung wissenschaftlicher Vorhaben (...) vom Statistischen Bundesamt und den statistischen Ämtern der Länder Einzelangaben an Hochschulen oder sonstige Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung übermittelt werden, wenn die Einzelangaben nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft zugeordnet werden können (...)”*.

Da §16(6) BStatG nur auf den ‚Grad‘ der Anonymisierung (faktische Anonymität) Bezug nimmt, sieht die Zensuskommission bei einer Datenweitergabe, die sich ausschließlich auf §16(6) BStatG stützt, insofern Probleme¹ als

- (1) der Zeitpunkt des Zugangs bzw. der Bereitstellung von Zensus-Mikrodaten völlig offen ist;
- (2) unklar ist, welche Zensus-Erhebungsmerkmale für die Forschung verfügbar sein werden;

¹ Die konkrete Umsetzung der faktischen Anonymisierung sowie der technisch und organisatorischen Randbedingungen der Datenbereitstellung werden im Folgenden ausgeklammert, da die Zensuskommission bereits empfohlen hat, hierfür Mittel für ein entsprechendes Forschungsprojekt bereitzustellen (Ergebnisniederschrift der Sitzung der Zensuskommission am 22. Januar 2009)

(3) das Risiko besteht, dass – ohne eine spezifische Wissenschaftsklausel im ZensG 2011 – letztendlich nur das Zensus-,Endprodukt' für Forschungsvorhaben verfügbar sein wird.

Die Beantragung und der Aufbau von thematischen und methodischen Zensus-Forschungsprojekten (z.B. bei der DFG) ist nur möglich, wenn der zur Verfügung stehende Merkmalskatalog zumindest in den Grundzügen sehr frühzeitig klar ist. Die Zensus-Kommission erachtet es in diesem Zusammenhang gleichfalls als wichtig, dass der Datenzugang zu einem möglichst frühen Zeitpunkt gewährt wird, optimalerweise nach Abschluss der Datenaufbereitungsarbeiten.

Der Zensus dient auch der Übermittlung von sozialstrukturellen Grunddaten an Eurostat. Zu diesem Zweck müssen Originalmerkmale z.T. entsprechend der Vorgaben von Eurostat in international harmonisierte Merkmale (z.B. ISCO, ISCED, ILO-Erwerbsstatus) transformiert werden. Diese von Eurostat geforderte Harmonisierung ist zentral für den europäischen und internationalen Vergleich, allerdings wird sie den nationalen Randbedingungen nicht immer gerecht, d. h. einige der nationalen Originalmerkmale werden aussagekräftiger als die EU-harmonisierten Daten sein. Für die Untersuchung einer Vielzahl von Forschungsfragen ist es aber unabdingbar, dass nicht nur die harmonisierten ‚Exportvariablen‘, sondern weiterhin auch die Originalmerkmale verfügbar sind. Nur so bleibt die Analyseflexibilität der Daten erhalten.

Bedingt durch die vorrangige Zielsetzung des Zensus ‚Festlegung gerichtsfester Einwohnerzahlen‘ wird die Datenqualität von Seiten des Gesetzgebers anhand des Kriteriums einer möglichst präzisen Schätzung der Bevölkerungszahl gemessen. Aus der wissenschaftlichen Perspektive ergibt sich durch das Modell des registergestützten Zensus eine ergänzende Sichtweise von Datenqualität. Denn die Kombination von Registerdaten und Primärerhebungen (Gebäude- und Wohnungszählung, Haushaltsstichprobe; Erhebung in Sonderbereichen) liefert nicht nur Erkenntnisse hinsichtlich Über- und Unterfassung, sondern es bestehen auch partielle Überschneidungen in

den Erhebungsmerkmalen von Registerdaten und Primärerhebungen (z.B. Ausbildung, Schulbildung, ausgeübter Beruf, Wirtschaftszweig, Familienstand, Staatsangehörigkeiten).² Über den systematischen Vergleich dieser Überschneidungsmerkmale lassen sich weitergehende Erkenntnisse über das Ausmaß und die Richtung von Abweichungen zwischen Register- und Primärdaten gewinnen. Diese sind sowohl für zukünftige registergestützte Zensen wie auch für die Survey-Methodenforschung von grundsätzlichem Interesse, da sich hieraus Anhaltspunkte für eine verbesserte Datenerhebung sowohl bei Registern wie auch Primärerhebungen ergeben.

Ein **typisches** und **aktuelles Beispiel**, für welche die Zensusdaten eine hervorragend geeignete Datenbasis bieten würden, wäre etwa der Vergleich des von Seiten der Arbeitgeber an die Bundesagentur für Arbeit gemeldeten ‚Ausbildungsniveaus‘ der Beschäftigten mit den Selbstauskünften zur Schul- und Berufsausbildung aus der Haushaltsstichprobe. Dies insbesondere deshalb, weil den Registerdaten der BA in den empirischen Wirtschaftswissenschaften eine herausgehobene Bedeutung zukommt auch in Hinblick auf politische Planung, die Qualität des Merkmals ‚Ausbildungsniveaus‘ jedoch als durchaus verbesserungsfähig eingeschätzt wird.³ Eine unabdingbare Voraussetzung für diese Art der Methodenforschung ist jedoch, dass in den Zensusdaten (bzw. in einer Arbeitsversion) die vorhandenen Überschneidungsmerkmale erhalten bleiben und der Wissenschaft auch zugänglich sind. Gleichfalls von erhöhten

² **Exkurs zu §7 Absatz 1 Satz ZensG 2011:** „Die Feststellung umfasst nicht die Berichtigung der aus den Melderegistern übernommenen Angaben zur Person, wie die Angabe ‚üblicher Aufenthaltsort‘ “. Nach dieser Festlegung haben die in Melderegistern vorliegenden Informationen Vorrang vor den primär erhobenen Angaben der Haushaltsstichprobe. Soweit sich dies auf das Merkmal üblicher Aufenthaltsort bezieht, ist der Vorrang des Melderegisters nachvollziehbar. Da die Erhebungsbeauftragten einen verkürzten Melderegisterauszug (Namen, Geschlecht, Tag der Geburt, Staatsangehörigkeiten der Befragungsperson) erhalten, sollten die Abweichungen zwischen Melderegister und Primärerhebung eher gering ausfallen. Was aber ist, wenn offensichtliche Abweichungen (z.B. Geburtsjahr) auftreten? Welche Angabe hat dann Vorrang, bzw. werden solche Abweichungen in einem Zwischenschritt (Zusammengeführte Daten nach §9) dokumentiert und stehen für Forschungszwecke zur Verfügung?

³ vgl. u.a.: Fitzenberger, Bernd, Aderonke Osikominu und Robert Völter (2006), Imputation Rules to Improve the Education Variable in the IAB Employment Subsample, in: **Schmollers Jahrbuch** Bd. 126, Heft 3.

Interesse ist auch, dass die methodischen Variablen für die Haushaltegenerierung in den Daten verfügbar bleiben.

Bei den oben angeführten Punkten handelt es sich um Detailfragen, die sicherlich nicht Gegenstand gesetzlicher Regelungen sein können und sollten. Sie wurden hier nur beispielhaft skizziert, um zu verdeutlichen, dass das Interesse der Forschung an der Nutzung dieser Daten nicht auf das (anonymisierte) Zensus-,Endprodukt' beschränkt ist. Vielmehr ergibt sich das eigentliche methodisch und inhaltliche Forschungspotenzial dieser Daten (im Vergleich zu anderen Datenquellen) erst aus der Kombination der verschiedenen Erhebungsteile und ihrer partiellen Überschneidungen, weshalb diese zusammengeführten Daten (in anonymisierter Form) grundsätzlich für Forschungsvorhaben zur Verfügung stehen sollten.