

**Beitrag von Destatis zur Beantwortung des Schreibens „Stellungnahme Zensus 2011“ der Stadt Magdeburg vom 3. Juli 2013**

Zu 1.

Die Hochrechnung der Einwohnerzahl erfolgt rein Design-basiert, d.h. die Einwohnerzahl wird als deterministische Größe angesehen, und man betrachtet nur die Wahrscheinlichkeit, dass eine Anschrift in die Stichprobe gelangt. Daher ist man – anders als bei üblichen Regressionsanalysen – von Modellannahmen bei der Regression unabhängig, und der relative Standardfehler ist dann die entscheidende Größe zur Beurteilung der Qualität. In der Formel für die Schätzung dieses Fehlers spielt die Varianz der Residuen die zentrale Rolle. Daher beeinflusst das Bestimmtheitsmaß den Zufallsfehler mittelbar, d.h. je höher das Bestimmtheitsmaß, desto kleiner ist der Zufallsfehler bei ansonsten gleichen Strukturen. Das Bestimmtheitsmaß der hinter der Schätzung stehenden Regression ist für Magdeburg 0,992. Nicht alle Regressionsparameter der exogenen Variablen dürften signifikant sein – das spielt aber aus dem oben erwähnten Grund auch keine Rolle. Zwar nützen die nicht signifikanten Variablen für die Einwohnerzahl nichts, aus Gründen der Kohärenz mit untergliederten Ergebnissen sind sie jedoch enthalten. Außerdem wenden wir aus Gründen der Gleichbehandlung für alle Gemeinden das gleiche Modell an.

Zu 2.

Die zitierte Formel mit dem Index d für einen Domain („Nachweisungsposition“) ist in dem wissenschaftlichen Gutachten von Münnich et al. bewusst zunächst allgemein gehalten. Für die Einwohnerzahlen der großen Gemeinden ist d daher die Gemeinde, d.h. die Schätzung erfolgt separat in jeder Gemeinde. Es war untersucht worden, die Schätzung gruppiert durchzuführen, d.h. die Regressionsparameter auf einer übergeordneten Ebene zu schätzen. In den Empfehlungen des Stichprobenforschungsprojekts haben die Auftragnehmer präzisiert, welche Gruppen gewählt werden sollten: „Als Gruppen eignen sich SMPs (Sampling Points), Kreise bzw. Bundesländer“. Die statistischen Ämter des Bundes und der Länder haben sich für die Umsetzung der Option „SMPs“ entschieden, um in die Berechnung der Einwohnerzahlen einer Gemeinde auch nur Informationen aus der betroffenen Gemeinde selbst einfließen zu lassen.

Ergänzend waren Münnich et al. gebeten worden zu prüfen, ob das Vorhandensein von Nullanschriften (= unbemeldete Anschriften, d.h. Anschriften von Gebäuden mit Wohnraum, zu der im Melderegister kein Personeneintrag vorhanden ist) in der Stichprobe in Verbindung mit den realen Verteilungen von Karteileichen und Fehlbeständen aus der Haushaltstichprobe des Zensus eine Nachjustierung des im Forschungsprojekt empfohlenen Hochrechnungsverfahrens erforderlich macht. Die Prüfung kann zu folgendem Ergebnis: "Als geeigneter Schätzer für die Anwendung auf der realisierten Stichprobe wird weiterhin der SMP-SEP GREG-Schätzer ohne Ausreißerbehandlung und ohne gesonderte Behandlung der Nullanschriften empfohlen." Es hatte sich gezeigt, dass die Wahl der SMPs als Gruppen ohne Ausreißerbereinigung und ohne Herausnahme von Nullanschriften die beste Wahl bei der Umsetzung der GREG-Schätzung ist, da hier die niedrigsten Standardfehler resultierten.

Zu 3.

Anschriften ohne gemeldete Personen sind Teil der Schicht 1 (unterste Anschriftengrößenklasse). Eine Einbeziehung unbemeldeter Anschriften, die nach §14 ZensusG positiv auf Wohnraum geprüft worden waren, in die Auswahlgesamtheit der Zensus-Stichprobe war methodisch zwingend, da andernfalls durch Vernachlässigung reiner Fehlbestandsanschriften eine systematische Untererfassung von Register-Fehlbeständen eingetreten wäre. In Magdeburg hatten 7,2% aller Stichprobenanschriften zum Stichtag keine gemeldeten Personen; in diesen wurden keine Fehlbestände beobachtet.

Die Aufteilung des Gesamtstichprobenumfangs auf die einzelnen SMPs und Anschriftengrößenklassen erfolgte nach einem Optimierungsalgorithmus, mit Hilfe dessen unter der Randbedingung, dass die Auswahlätze nicht zu groß und nicht zu klein werden, der Gesamtfehler deutschlandweit minimiert wurde. Der realisierte Auswahlatz für eine konkrete Anschriftengrößenklasse floss in die Berechnung der Hochrechnungsfaktoren ein. Bei diesem bewusst disproportionalen Stichprobendesign kann daher u.E. nicht von einer Über- oder Unterrepräsentierung einzelner Schichten gesprochen werden. Im Übrigen beträgt der Anteil der Varianz der Schicht 8 an der gesamten Fehlervarianz (quadrierter absoluter Standardfehler) nur 6,2%.

Zu 4.

Die Stadt Magdeburg sieht es aufgrund der Ergebnisse der wissenschaftlichen Studie (Regressionsanalyse) von Hoppe und Spandel (2013) als erwiesen an, dass die Unterschiede zwischen den kleinen Gemeinden (unter 10 000 Einwohner) und den großen Gemeinden (10 000 und mehr Einwohner) bezüglich der relativen Abweichungen der mit dem Zensus 2011 ermittelten neuen Einwohnerzahlen von den alten Einwohnerzahlen der bisherigen Bevölkerungsfortschreibung auf die unterschiedlichen Berechnungsmethoden der Einwohnerzahlen beim Zensus 2011 zurückzuführen seien. Dieser Schlussfolgerung aus den Ergebnissen der Regressionsanalyse kann sich das Statistische Bundesamt aus folgenden Gründen nicht anschließen.

- Die Einwohnerzahlen aus der bisherigen Bevölkerungsfortschreibung dürfen nicht mit den aus den Melderegistern für den Zensusstichtag ermittelten Einwohnerzahlen gleichgesetzt werden. Mit den Ergebnissen der Haushaltstichprobe bzw. der Befragung zur Klärung von Unstimmigkeiten sollen nämlich die aus den Melderegistern ermittelten Einwohnerzahlen um Überfassungen und Untererfassungen korrigiert werden und nicht die Einwohnerzahlen gemäß Bevölkerungsfortschreibung. Damit ist die Wahl der endogenen Variablen für die zitierten Regressionsanalysen nicht sachadäquat erfolgt. Die relative Abweichung der neuen von den bisherigen Einwohnerzahlen kann nämlich von Faktoren stark beeinflusst sein, die nichts mit der Qualität der Melderegister zu tun haben. So können z.B. die Einwohnerzahlen aus der bisherigen Bevölkerungsfortschreibung überhöht sein, wenn die Meldebehörden nicht alle Abmeldungen von Amts wegen an die Statistischen Ämter der Länder weitergeleitet haben. Bei nicht wenigen Gemeinden dürften auch schon die Basiszahlen der bisherigen Bevölkerungsfortschreibung (Ergebnisse der Volkszählung 1987 bzw. das zentrale Einwohnermelderegister der DDR, Stand Oktober 1990) überhöht gewesen sein.
- Fraglich ist ebenfalls, ob die erklärende Variable „kleineGemeinde/große Gemeinde“ ein guter Proxy für die unterschiedliche Berechnungsmethode der Einwohnerzahl ist. Zu berücksichtigen ist hier, dass man mit der erklärenden Variablen „kleine Gemeinde/große Gemeinde“ unbeabsichtigt andere Effekte als die Berechnungsmethodik messen kann, die mit der Gemeindegröße korreliert sind, z.B. Stadt-Land-Unterschiede. Die Unterschiede bei den Abweichungen der alten von den neuen Einwohnerzahlen müssen also nicht zwangsläufig ihre Ursachen in der unterschiedlichen Berechnungsmethode haben.
- Das Bestimmtheitsmaß der Regression ist zu gering, als dass der (zwar signifikante) Regressionskoeffizient für die Dummy-Variable "Methodik" einen starken Einfluss dieser erklärenden Variablen auf die endogene Variable belegen würde; vielmehr ist der Einfluss auf Grund des sehr niedrigen Bestimmtheitsmaßes schwach.

- Die vorgestellten Zahlen zeigen zwar, dass die Einwohnerzahlen aus der bisherigen Bevölkerungsfortschreibung keinen signifikanten LINEAREN Einfluss auf die relative Änderung dieser fortgeschriebenen EWZ durch den Zensus hat, aber die Regressionsmodelle machen keinerlei Aussage darüber, ob es nicht vielleicht einen starken signifikanten NICHT-LINEAREN Zusammenhang zwischen EWZ und EWZ-Änderung gibt.