

Dr. Barbara Wawrzyniak

Small Area Methoden: Neue Regeln zur Berechnung regionaler Markt- und Potenzialdaten.

Es besteht ein zunehmender Bedarf an regionalisierten Daten, die feinräumigere und somit präzisere Ergebnisse für Gebietsplanungen, Potenzial- oder Marktanalysen liefern. Als feinräumig werden dabei alle Einheiten unterhalb der Gemeindeebene angesehen. Besonders effizient und verlässlich sind hier die statistischen Verfahren der „Small Area Methods“ (SAM).

Die Small-Area-Methoden (SAM) beschäftigen sich im Wesentlichen mit der Entwicklung und Verbesserung statistischer Verfahren, auch Small Area Estimates oder Small Area Statistics genannt, zur Schätzung kleinräumiger Daten, bei denen die Regionen nur sehr wenige oder gar fehlende (Befragungs-) Werte aufweisen. Im Zentrum steht dabei das Regionalisieren, also das „Herunterrechnen“ von für größere Regionen vorliegenden aggregierten Daten in feinere geographische Einheiten (Top-down-Verfahren).

Spezialisten wie infas 360 haben dafür eigene, auf amtlicher Basis entwickelte geographische Einheiten, die Small Statistical Areas, die durch eine Vielzahl von Geo- und Marktdaten (Lage, Typ, Größe und Baujahr von Gebäuden, Wohnungs- oder Haushaltsanzahl, Alter der Haushaltsvorstände, KFZ-Daten, Kaufkraft, Lebensstil, usw.) präzise charakterisiert sind. Diese Small Statistical Areas lassen sich wiederum zu jedem beliebigen Gebiet aggregieren und passen in alle postalisch-amtlichen Systeme. Vorhandene Echtdateien für bestimmte geographische Einheiten können so anderen geographisch und inhaltlich ähnlichen Einheiten zugeordnet, also „geschätzt“ werden. Umgekehrt lassen sich Gebiete, für die bestimmte Echtdateien vorliegen, durch die feinräumige Beschreibung der Small Statistical Areas typologisieren, was etwa für die Ursachenforschung eines gesellschaftlichen oder wirtschaftlichen Phänomens innerhalb eines Gebietes hochinteressant ist.

Mit Small-Area-Methoden (SAM) können für jede geographische Ebene Mittelwerte, Anteile, Proportionen oder Raten geschätzt werden. Besonders interessant sind Small-Area-Methoden für die Regionalisierung amtlicher Daten wie z.B. Einkommen oder Arbeitslosenzahlen. Diese liegen meistens nur auf eher groben räumlichen Ebenen der Kreise oder Gemeinden vor. Mit Hilfe von Befragungsdaten (dabei sind nur geringe Fallzahlen erforderlich) lassen sich diese Werte auf kleinere Einheiten, wie den Ortsteil oder den Siedlungsblock rechnen. Das Schätzen von Marktanteilen auf kleinräumiger Ebene kann ebenfalls auf diese Weise verlässlich gerechnet werden.

Die Vorteile von SAM gegenüber herkömmlichen Schätzmodellen bestehen darin, dass in das Modell Informationen aus **übergeordneten** Raumebenen ebenso einfließen wie Informationen aus **ähnlichen** Regionen. Das führt zu sehr genauen Angaben, selbst wenn die Regionen nur kleine oder gar fehlende Fallzahlen aufweisen. Voraussetzung der Raumebenen ist eine hierarchisch überschneidungsfreie Struktur (**„nested data“**)

Dabei wird angenommen, dass bestimmte Zusammenhänge zwischen Variablen nicht nur in einem Gebiet Gültigkeit besitzen, sondern generell in der Population auf ähnliche Weise vorliegen. Somit können Informationen aller oder zumindest vieler Regionen zu einer Stabilisierung der Schätzungen in den einzelnen Regionen beitragen. Prof. Calyampudi Radhakrishna Rao (2004) spricht dabei von „borrow strength“ from related areas to find indirect estimates that increase the effective sample size

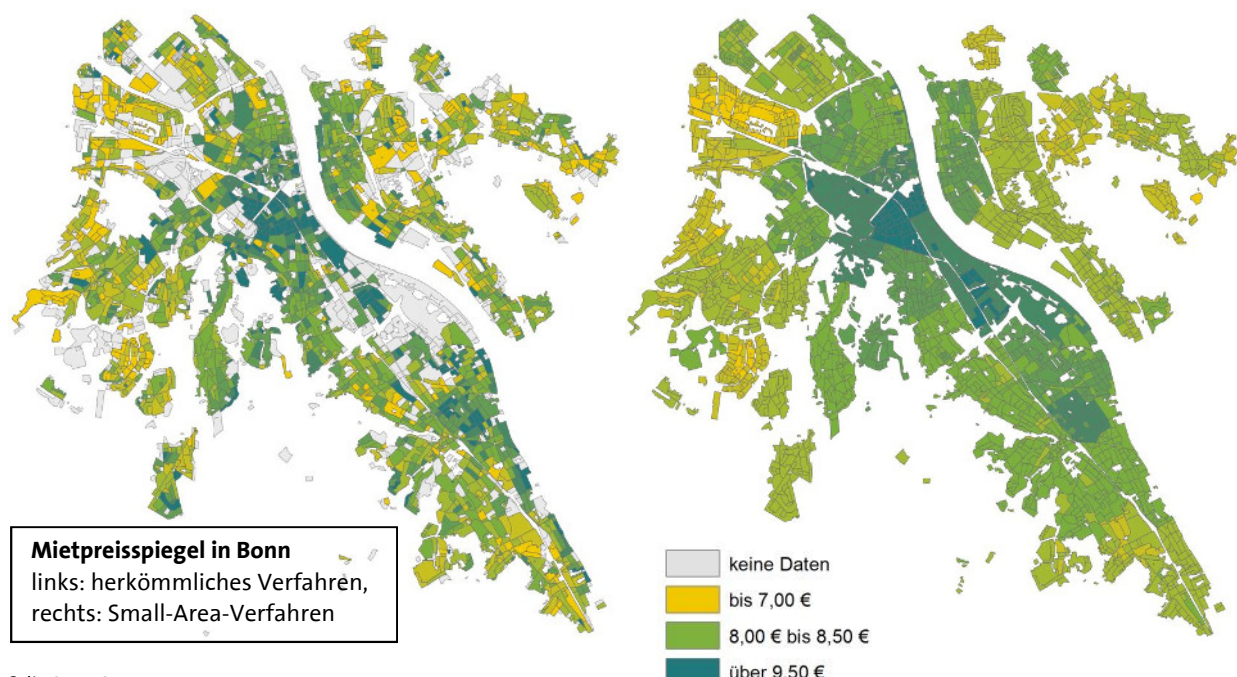
and thus increase the precision.“ Wichtige Zusatzinformationen liefern dabei sogenannte Hilfsvariablen für die jeweilige Region, die das Modell erheblich verbessern können. Mögliche Hilfsinformationen sind z.B. Merkmalsausprägungen aus der Vergangenheit, Werte aus benachbarten oder übergeordneten Subpopulationen und/oder Werte von Hilfsvariablen, die stark mit dem interessierenden Merkmal korrelieren. Es wird generell versucht, Gemeinsamkeiten und Unterschiede verschiedener Gebiete in einem Modell abzubilden, und dann dieses Modell etwa für die Vorhersage der Populations- oder andere Kennwerte zu verwenden.

In den USA sind die Small-Area-Methoden bereits weit verbreitet. So wird dort über dieses Modell beispielsweise die regionale Armut in der Erhebung SAIPE abgeleitet. In Europa halten die Methoden erst in den vergangenen Jahren Einzug in amtliche Datenbestände und Vorgehensmodelle. Viele Staaten sind bei Befragungen bereits von der klassischen Vollerhebung abgerückt und zu einer kostengünstigeren Stichprobenerhebung übergegangen, die durch die Small-Area-Methoden unterstützt werden. Der Einsatz von SAM in den deutschen Zensus wird ebenfalls diskutiert.

Der Einsatz dieser statistischen Verfahren ermöglicht die nachvollziehbaren Übertragungen von Echtdaten und Kennziffern auf jedes beliebige Gebiet. Bereits vorhandene Befragungsdaten lassen sich mittels Small Area Statistics sehr leicht nachträglich regionalisieren. So werden geographische und Zielgruppen-Analysen jeder Art kostengünstig, flächendeckend, standardisiert und nachvollziehbar möglich. Über die staatlich-amtlichen Anwendungen hinaus wird das Modell deshalb auch immer interessanter für den wirtschaftlichen Sektor.

Anwendungsbeispiel Mietpreisspiegel

Ein konkreter Anwendungsfall, der gerade auch durch die gesetzliche Mietpreisbremse an Aktualität gewinnt, ist die Berechnung von durchschnittlichen Mietpreisen für einzelne Ortsteile oder Siedlungsblöcke. Beim herkömmlichen Vorgehen würde man die erhobenen Mietpreise als Durchschnittswerte berechnen, was bei Vorliegen von ausreichend hohen Fallzahlen durchaus seine Berechtigung hat.



Auf kleinräumiger Ebene tritt aber häufig der Fall auf, dass nicht genügend Werte für eine verlässliche Berechnung vorliegen. Dann wird mit Hilfe der SAS die Schätzung der Mietpreise für alle Siedlungsblöcke ermöglicht, in dem nicht nur die Einzelmieten in das Modell einfließen, sondern auch Informationen zu den Siedlungsblöcken selber und der darüber liegenden Ebene, dem Ortsteil.

Für diese Herangehensweise bildet das hierbei verwendete Postalisch-Amtliche-Gliederungssystem (PAGS) von infas 360 die idealen Voraussetzungen. PAGS basiert auf allen flächendeckend verfügbaren amtlichen Gebäudedaten und beinhaltet sämtliche postalisch-amtlichen Strukturen.

PAGS berücksichtigt Deutschlands Gemeinden und Postortsteile und weist alle bebauten (Siedlungen) und nicht bebauten Flächen aus. Innerhalb dieser werden Quartiere nach Gewerbe, Misch- oder Wohnnutzung unterteilt, wobei eine homogene Bebauungsstruktur maßgeblich ist. Damit bildet PAGS erstmals die siedlungsstrukturelle Realität des Raums ab. Dieses völlig neue Gliederungssystem ist die Basis aller präzisen und amtlich-postalisch nachvollziehbaren nachgelagerten Analysen und Datenberechnungen.

Autorin:

Dr. Barbara Wawrzyniak
infas 360 GmbH, Projektleiterin Customer Analytics

Studium der Sozialwissenschaften mit Schwerpunkt Familiensoziologie und Methoden der empirischen Sozialforschung sowie der Medien- und Kulturwissenschaften an der Universität Düsseldorf. Auslandssemester an der Jagiellonen-Universität Krakau/Polen. Promotion an der Universität Duisburg-Essen mit Schwerpunkt Lebensverläufe und Längsschnittanalysen. Von 2010 - 2013 Projektkoordinatorin an der Universität zu Köln sowie lokale Datenkoordinatorin für die Universität Bamberg. Langjährige Tätigkeit in Jour