

Neuorientierung der Methoden-Ausbildung

Ein kurzer Problemaufriss¹

Andreas Diekmann

Wie in vielen Disziplinen gibt es heute eine teilweise erhebliche Diskrepanz zwischen Lehrbuchwissen/Inhalten von Methodenkursen und dem Stand der Forschung. Man könnte mit Ogburn von einem »cultural lag« sprechen; einer Lücke, die auch durch den rasanten Fortschritt von Digitalisierung und Technologieentwicklung bedingt ist. Das bisherige Programm von Methodenkursen muss modernisiert und an diese Entwicklungen angepasst werden. Wenn das nicht geschieht, wird die Soziologieausbildung gegenüber anderen Sozialwissenschaften ins Hintertreffen geraten. Neuentwicklungen gibt es im Bereich der *Datenquellen*, dem *Design von Erhebungen*, der *Datenanalyse* und der *Archivierung der Daten* – Stichwort »Open Science«. Dieser kurze Problemaufriss soll einen Anstoß zur Reform der Methodenausbildung geben.

Neue Daten

Beruhte der Großteil standardisierter Forschung auf Surveydaten, so sind durch Digitalisierung zahlreiche neue Datenquellen hinzugekommen. Dazu zählen Sensordaten in unterschiedlichen Bereichen, Texte im WWW, aus denen mit Webscraping strukturierte Daten gewonnen werden können und georeferenzierte Daten, mit denen kleinräumig zu verschiedenen sozialwissenschaftlichen Aspekten – sowohl für explorative als auch und besonders für theorieorientierte Studien – ein großes Potential neuer Informationen vorliegt. Oft können diese Daten auch mit Surveydaten verlinkt werden.

¹ Dieser Text war eine Vorlage für das Konzil der DGS und wurde in der Sitzung am 29. Oktober 2022 in Bielefeld zur Diskussion gestellt.

Einige Beispiele: (1) Über die Auswirkungen des 9-Euro-Tickets auf veränderte Mobilitätsmuster informieren unter anderem Mobilfunkdaten. (2) Per Webscraping wurden in einer Studie mehrere Hunderttausend eBay-Auktionen analysiert, um neue Erkenntnisse über Reputationseffekte bei wirtschaftlichen Transaktionen zu gewinnen. (3) In einer soziologischen Umweltstudie wurden Surveydaten aus der Befragung von Haushalten mit georeferenzierten Registerdaten von Lärm und Luftverschmutzung verknüpft, so dass zum Beispiel die Stärke der Abhängigkeit der Umweltbelastung vom Sozialstatus und anderen Merkmalen des Haushalts ermittelt werden kann. (4) Früher wurden zum Beispiel Hemmnisse der Erwerbstätigkeit von Frauen ausschließlich per Surveystudien untersucht. Heute kann man mit georeferenzierten Daten den räumlichen Kontext beschreiben, die Distanz zu Kinderkrippen, Schulen, Einkaufsmöglichkeiten, ÖV-Haltestellen, Naherholung, Parks und vieles andere mehr. Diese Daten können zudem mit Befragungsdaten anonymisiert verknüpft werden. Das spart nicht nur Zeit bei Surveys, sondern ermöglicht auch validere Aussagen als bei ausschließlich selbstberichteten Informationen. Ferner können die räumlichen Daten in Geoinformationssystemen (GIS) ausgewertet und grafisch dargestellt werden. Um diese Datenquellen zu erschließen und auszuwerten, bedarf es aber auch entsprechender Kompetenzen, die in der Ausbildung bisher nicht vermittelt wurden.

Erhebungsdesigns

Querschnitts-Korrelationsanalysen führen häufig in die Irre, wenn es um die Aufklärung kausaler Zusammenhänge geht. Seit einigen Jahren werden auch in den Sozialwissenschaften vermehrt *randomized controlled trials* durchgeführt (RCT-Studien). Für den innovativen Einsatz dieser Methodik in der Armuts-, Gesundheits- und Entwicklungsforschung haben kürzlich Abhijit Banerjee und Esther Duflo den Nobelpreis erhalten. Nicht immer sind RCT-Designs realisierbar, aber es gibt oft Näherungen, die in Verbindung mit entsprechenden Auswertungsmethoden valide Erkenntnisse zur Wirkung (oder Nicht-Wirksamkeit) von Maßnahmen liefern (zum Beispiel Diff-in-Diff-Design). Die Expertenkommission zur Wirksamkeit von Maßnahmen in der Corona-Krise hat festgehalten, dass man außer der Wirksamkeit von Masken eigentlich überhaupt nicht weiß, in welchem Ausmaß die zahlreichen Maßnahmen wie Lockdown oder Schulschließung die Ausbreitung der Epidemie gebremst haben. Mehr als kleine Befragungsstudien (zum Beispiel die COSMO-Studien)

gab es in Deutschland in den Sozialwissenschaften nicht. Aussagekräftige Erhebungen und Designs wären in vielen Bereichen einsetzbar. Auch diese Kompetenzen sollten in der Methodenausbildung vermittelt werden.

Datenanalyse

Früher hat man viel zu optimistisch geglaubt, dass Querschnitts-Surveydaten Aufschluss über kausale Effekte geben können, wenn nur für möglichst viele andere Merkmale »kontrolliert« wird. *Overcontrol* und unbeobachtete Heterogenität kann aber zu krassen Fehlschätzungen führen. Beispiel: Regressions-schätzungen mit »Kontrollvariablen« zeigen, dass Kohabitation vor der Eheschließung das Scheidungsrisiko erhöht; der tatsächliche kausale Effekt ist aber genau umgekehrt. In der Gesundheitsforschung wurde lange Zeit von einer U-Kurve zwischen Alkoholkonsum und Mortalität gesprochen. Abstinenz wäre demnach ungesund, mäßiger Konsum von Alkohol gesünder, viel Alkohol natürlich ungesund. Leider ist es nicht so. Heute weiß man, dass die Abstinenzler in den Studien eine teilweise selektive Gruppe von Personen waren, die aus Krankheitsgründen keinen Alkohol konsumierten oder früher Alkoholiker waren. Tatsächlich steigt das Risiko monoton mit dem Konsum. Inzwischen unterscheidet man auch genauer bei multivariablen Analysen zwischen verschiedenen Arten von potentiell unabhängigen Variablen: *Confounders*, *Mediators*, *Colliders*. Hier zeigt sich übrigens auch wieder die Bedeutung theoriegeleiteter Forschung. Denn die Rolle, die die Variablen einnehmen, hängt auch von theoretischen Vorüberlegungen ab. *Confounders* sollen in Regressionsschätzungen aufgenommen werden, *mediators* und *colliders* nicht, wenn das Ziel die näherungsweise Schätzung kausaler Einflüsse ist. Bei Querschnittsdaten bleibt natürlich das Problem, dass man gar nicht alle *confounders* kennt (unbeobachtete Heterogenität). Seit einigen Jahren (in der statistischen Literatur einigen Jahrzehnten) wird diesen Problemen aber in wachsendem Maße in den Sozialwissenschaften Rechnung getragen. Und zwar mittels Paneldaten und *Fixed-Effects*-Regression, wodurch (zeitkonstante) unbeobachtete Heterogenität quasi ausgeblendet werden kann. Die Anwendung dieser Modelle ist ein ganz wesentlicher Fortschritt! Es ist zwar nicht der Goldstandard, aber auf diese Weise können kausale Wirkungen anhand nicht-experimenteller Daten relativ verlässlich ermittelt werden. In der Soziologie werden diese Verfahren auch häufig angewandt – ich staune über die vielen professionellen Arbeiten vieler unserer Nachwuchswissenschaft-

lerinnen und Nachwuchswissenschaftler. Aber die Varianz der Kompetenzen ist in unserem Fach sehr groß! Deshalb gehören Grundkenntnisse dieser Methoden in die Standard-Methodenausbildung!

Noch ein Punkt. In den letzten Jahren wächst die Kritik an der Signifikanzstatistik. Ich würde nicht so weit gehen und einer Abschaffung das Wort reden, wie das bereits Fachzeitschriften praktizieren. Die ritualisierte Anwendung ist aber höchst fragwürdig – Ziliak und McCloskey kritisieren schon 2008 »The Cult of Statistical Significance«. Der Medizinstatistiker Ioannidis hat 2005 in seinem vielzitierten Artikel »Why most published research findings are false« (über 11.600 Zitationen bei Google Scholar!) auf die Überschätzung der Aussagekraft der Tests hingewiesen. Psychologische Fachzeitschriften setzen derzeit neue Standards, welche Parameter bei dem Einsatz von Signifikanztests berichtet werden sollen (Sternchen allein genügen sicher nicht!). Die kritische Einschätzung von Signifikanztests muss auch in der Ausbildung stärker zur Sprache kommen!

Open Science

Nicht zuletzt durch Fälschungsskandale und Nicht-Reproduzierbarkeit hat die Open-Science-Initiative an Unterstützung gewonnen. Dazu gehört vor allem die niedrighschwellige Verfügbarmachung dokumentierter Forschungsdaten in einem Datenarchiv; des Weiteren die Präregistrierung von Hypothesen und die Setzung von Anreizen für (sorgfältige) Replikationsstudien. Präregistrierung ist natürlich nur möglich bei theoriegeleiteter Forschung, nicht bei explorativen Studien (die zweifellos sehr sinnvoll sein können). Für Replikationen gilt: Jede gut gemachte und methodisch solide Replikationsstudie ist ein Beitrag zum Kollektivgut der Verlässlichkeit wissenschaftlicher Erkenntnisse. Weiteres zum Open-Science-Projekt findet man in dem von Brian Nosek mitgegründeten »Center for Open Science« (www.cos.io). Wichtig ist, dass auch in der Ausbildung für diese Themen sensibilisiert wird!