

Analysepotenziale sozialwissenschaftlicher Forschungsdaten

Workshop 24. bis 26. September 2012
an der Universität Bielefeld



Schon länger besteht in den Sozial- und Wirtschaftswissenschaften Einigkeit darüber, dass viele Forschungsfragen nur im Längsschnitt adäquat untersucht werden können. Längsschnittliche Forschungsdaten, die auf einem Panel- oder Trenddesign beruhen, werden deshalb immer häufiger genutzt. Bereits die bestehenden Längsschnittstudien setzen aufgrund ihrer Datenstruktur bestimmte Kenntnisse und Erfahrungen im Datenhandling voraus, die in der universitären Lehre nur begrenzt vermittelt werden (können).

Die dadurch bestehenden Hürden beim sekundäranalytischen Zugang zu diesen Daten werden noch verstärkt, weil viele Forschungsdatenprojekte in den letzten Jahren die Möglichkeiten der Verknüpfung mit anderen Datenquellen erfolgreich nutzen. Die daraus entstehenden hochkomplexen Datenstrukturen bieten ein historisch einmaliges Analysepotential für die sozial- und wirtschaftswissenschaftliche Forschung, das sich aber nur dann erschließt, wenn man die Struktur dieser Daten genau kennt und zugleich über die praktischen Kompetenzen im Umgang mit derartigen Strukturen verfügt.

Neben der Zunahme der Komplexität der „klassischen“ Forschungsdaten ist die Entwicklung der empirisch arbeitenden Sozial- und Wirtschaftswissenschaften in den letzten Jahren vor allem auch durch die Erschließung neuer Datenbestände geprägt.

Dies gilt besonders für Daten über Betriebe, Unternehmen und andere Organisationen. Informationen über die Strukturen und das Handeln von Organisationen sind nicht nur für die Organisationsforschung im engeren Sinne, sondern zunehmend auch für andere Forschungsfelder wichtig z.B. der Arbeitsmarktforschung, der Konjunkturforschung oder der familien- oder ungleichheitssoziologischen Forschung. Diese Datenbestände finden in den Ausbildungsprogrammen der unterschiedlichen Fachdisziplinen bisher nur sehr eingeschränkte Beachtung.

Vor dem Hintergrund der beiden skizzierten Entwicklungen werden beim diesjährigen *data lab*-Workshop vom 24. bis 26. September 2012 an der Universität Bielefeld die Nutzungsmöglichkeiten des Sozio-ökonomischen Panels (SOEP, Paneldatensatz), des SHARE-Datensatzes (internationale, replikative Studie) sowie einiger ausgewählter Betriebs- und Organisationsdatensätze (IAB-Betriebspanel, AFiD der amtlichen Statistik sowie „Betriebe und Beschäftigung im Wandel“ [SFB Jena]) vorgestellt.

Die Grundidee des Workshops besteht darin, eine praxisorientierte Einführung in die Nutzung dieser Datensätze zu geben, d.h. neben der Vermittlung von Basisinformationen zum Design, Aufbau und den Nutzungspotentialen der Datensätze werden erste praktische Erfahrungen in der Nutzung durch intensive Übungen am Computer vermittelt. Die Vorstellung der jeweiligen Datensätze erfolgt durch Mitarbeiter der Forschungsdatenzentren der jeweiligen Datenproduzenten. Die Veranstaltung richtet sich insbesondere an Doktorand/innen und Nachwuchswissenschaftler/innen aus den Fächern Ökonomie, Politikwissenschaft, Psychologie und Soziologie.

Der Workshop besteht aus drei parallel angebotenen Veranstaltungslinien:

Stream 1: Einführung in die Benutzung des Sozio-ökonomischen Panel (SOEP)

Das SOEP ist eine seit 1984 laufende jährliche Wiederholungsbefragung in Deutschland, mit den Themenschwerpunkten Haushaltszusammensetzung, Erwerbs- und Familienbiographie, Einkommensverläufe, Gesundheit und Lebenszufriedenheit, Persönlichkeit und Kompetenzen.

Der Workshop bietet einen fundierten Einblick in das Arbeiten mit dem SOEP. So wird es einen umfassenden Überblick über Aufbau und Analysepotentiale des SOEP gegeben. Außerdem werden verschiedene methodische Aspekte (wie Stichprobenziehung, Gewichtung und Imputation) vorgestellt. Schließlich stehen die verschiedenen Möglichkeiten zur Durchführung intergenerationaler Analysen mit dem SOEP im Blickpunkt.

Stream 2: Einführung in die Benutzung von Survey of Health, Ageing and Retirement of Europe (SHARE) und Zusatzdaten

Der Survey of Health, Ageing and Retirement in Europe erhebt seit 2004 alle zwei Jahre Informationen über den Gesundheitszustand, den sozio-ökonomischen Status sowie die sozialen Netzwerke von mittlerweile über 80.000 Europäern ab 50 Jahren in 20 Ländern. Nach einem Einblick in die Analysepotenziale der ersten vier Wellen wird die Verknüpfung von SHARE DE mit Daten der deutschen Rentenversicherung vorgestellt. Im Anschluss daran stehen die retrospektive Lebensverlaufshebung SHARELIFE und die Verknüpfung mit Kontextdaten sowie Visualisierung mithilfe von Geo-Daten auf dem Übungsprogramm

Stream 3: Betriebs- und Unternehmensdaten (AFiD, IAB-Betriebspanel und „Betriebe und Beschäftigung im Wandel“)

Mit dem Projekt „*Amtliche Firmendaten für Deutschland (AFiD)*“ hat sich das FDZ der Statistischen Ämter der Länder das Ziel gesetzt, die bisher nur im Querschnitt und unverknüpft vorliegenden Einzeldaten der amtlichen Wirtschafts- und Umweltstatistiken in einem System verschiedener Paneldatensätze und zuspieldar Module zusammenzuführen. Somit können für einzelne Betriebe und Unternehmen die Informationen aus verschiedenen Statistiken zusammen genutzt werden. Auch Analysen zeitlicher und wirtschaftsübergreifender Entwicklungen werden erleichtert. Die Veranstaltung bietet einen Überblick über die möglichen Merkmals- und Statistikkombinationen. Nach einer kurzen methodischen Einführung soll dieser erste Eindruck mit einigen beispielhaften Fragestellungen zum AFiD-Panel Industriebetriebe, welches pro Welle die Angaben von ca. 50.000 Betrieben des produzierenden Gewerbes umfasst, in den Übungen vertieft werden.

Das *IAB-Betriebspanel* ist eine jährliche repräsentative Betriebsbefragung mit ca. 16.000 Betrieben aller Wirtschaftszweige und Größenklassen in Deutschland. Sie umfasst ein sich wiederholendes Standard-Fragenprogramm zu betrieblichen Bestimmungsgrößen der Beschäftigung und jährlich wechselnde Themenschwerpunkte. Im Rahmen der Übungen sind die Aufbereitung des Paneldatensatzes sowie ausgewählte Analysen geplant

Die ersten drei Wellen der im Rahmen des SFB 580 erhobenen Betriebsbefragung „*Betriebe und Beschäftigung im Wandel*“ stehen erst seit Kurzem im DSZ-BO für Sekundäranalysen zur Verfügung. Im Fokus der Erhebung stehen Struktur und Stabilität von Beschäftigungsverhältnissen.

Ablauf

Montag, 24 September 2012			
	Stream 1: SOEP	Stream 2: SHARE	Stream 3: Betriebs-/Unternehmensdaten
10.00-12.00	Propädeutikum in Stata für Teilnehmer/innen ohne Stata-Vorkenntnisse		
12.00-13.00	Mittagspause		
13.00-13.30	Begrüßung und Einführung		
13.30-15.00	Einführung SOEP Datenstruktur, Inhalte und Dokumentation	Einführung SHARE Datenstruktur und Dokumentation	Einführung Organisations- & Betriebsdaten Das IAB Betriebspanel (BP)
15.00-15.30	Pause		
15.30-17.00	Übung I: SOEP	Übung I: SHARE Datenmanagement	Übung I: IAB BP
17.00-17.30	Pause		
17.30-18.30	Fortsetzung Übung I: SOEP	Fortsetzung Übung I: SHARE	Fortsetzung Übung I: IAB BP
19.00-21.00	Vortrag: „Verfahren zur Längsschnittanalyse: Ein Überblick“ (Marco Gießelmann, DIW Berlin) Anschließend: Wine and Cheese		
Dienstag, 25. September 2012			
	Stream 1: SOEP	Stream 2: SHARE	Stream 3: Betriebs-/Unternehmensdaten
9.30-11.00	Das intergenerationale Design des SOEP: Verfügbare Informationen, Datenstrukturen und Anwendungsbeispiele	SHARE-RV: Verknüpfung mit den deutschen Rentenversicherungsdaten	Betriebs- und Unternehmensdaten der amtlichen Statistik AFID
11.00-11.30	Pause		
11.30-13.00	Übung II: SOEP	Übung II: SHARE RV Datenmanagement	Übung II: IAB BP
13.00-14.00	Mittagspause		

14.00-16.00	Übung III: SOEP	Fortsetzung Übung II: SHARE RV Versicherungsbiographien	Übung III: AFID
16.00-16.30	Pause		
16.30-18.30	Übung IV: SOEP	SHARELIFE Lebensverläufe und Kontextinformationen	Fortsetzung Übung III: AFID

Mittwoch, 26 September 2012			
	Stream 1: SOEP	Stream 2: SHARE	Stream 3: Betriebs-/Unternehmensdaten
9.00-10.30	Stichproben, Gewichtung und Hochrechnung im SOEP	Übung III: SHARELIFE Lebensverläufe im historischen Kontext	SFB580-B2-Betriebsbefragung
10.30-11.00	Pause		
11.00-12.15	Vortrag: „Die Imputation fehlender Werte“ (Prof. Dr. Cornelia Weins, Ruhr-Universität Bochum)		
12.15-13.00	Mittagspause		
13.00-15.00	Individuelle Beratung	Fortsetzung Übung III: SHARE und Geodaten – Visualisierung regionaler Muster	Übung I: SFB580-B2-Betriebsbefragung
15.00-15.15	Pause		
15.15-17.00	Individuelle Beratung	Individuelle Beratung	Übung II: SFB580-B2-Betriebsbefragung
17.00	Abschluss		

Teilnahme

Die Gebühr für die Teilnahme am Workshop beträgt 60 € (Doktorand/innen und Student/innen 35 €).

In der Teilnahmegebühr sind enthalten: Dokumentation des Workshops mit Präsentationen, Stata-Beispielprogrammen und Arbeitshilfen sowie Getränke und Obst in den Pausen und ein Wine&Cheese-Bufferet.

Anmeldung und weitere aktuelle Informationen

Online-Anmeldung (ab 12. Juli 2012) und weitere Infos unter:
<http://www.uni-bielefeld.de/soz/datalab2012/>

Vorkenntnisse

In den computerbezogenen Lehreinheiten wird die Statistik-Software Stata verwendet. Teilnehmer ohne Stata-Kenntnisse sollten deshalb die Einführungsveranstaltung am Montagmorgen besuchen.

Zielgruppe

Der Workshop richtet sich insbesondere an Doktorand/innen und Nachwuchswissenschaftler/innen,

- die für ihre theoretischen Fragestellungen geeignete Datensätze zur Sekundäranalyse kennen lernen wollen,
- und eine anwendungsbezogene Einführung in das Arbeiten mit komplexen Datensätzen mit individueller Betreuung suchen.

Veranstalter

Datalab2012 ist eine gemeinsame Veranstaltung des Ständigen Ausschusses Forschungsdateninfrastruktur (FDI) des Rat für Sozial- und Wirtschaftsdaten (RatSWD), des Arbeitsbereichs V der Fakultät für Soziologie der Universität Bielefeld und des Sonderforschungsbereich 882 „Von Heterogenitäten zu Ungleichheiten“ an der Universität Bielefeld in Kooperation mit den Forschungsdatenzentren SOEP, SHARE, der statistischen Landesämter und der BA im IAB sowie dem Datenservicezentrum für Betriebs- und Organisationsdaten.

Lokale Veranstalter

Arbeitsbereich V Sozialstruktur und Soziale Ungleichheit, Fakultät für Soziologie, Universität Bielefeld (Prof. Dr. Stefan Liebig, Prof. Dr. Martin Diewald)
Organisation: Dr. Christian Schmidt (Geschäftsführung SFB 882)

In Kooperation mit

Dr. Marco Gießelmann, Dr. Daniel Schnitzlein (SOEP/DIW Berlin)
Dr. Martina Brandt (SHARE MEA, MPISOC München)
Dr. Katharina Cramer, Matthias Klumpe, Michael Rößner (FDZ der statistischen Landesämter)
Dr. Susanne Kohaut und Dana Müller (IAB-Betriebspanel und FDZ der BA im IAB, IAB/Nürnberg)
Dr. Peter Jacobebbinghaus, Jelena Hohlweg (DSZ-BO)

