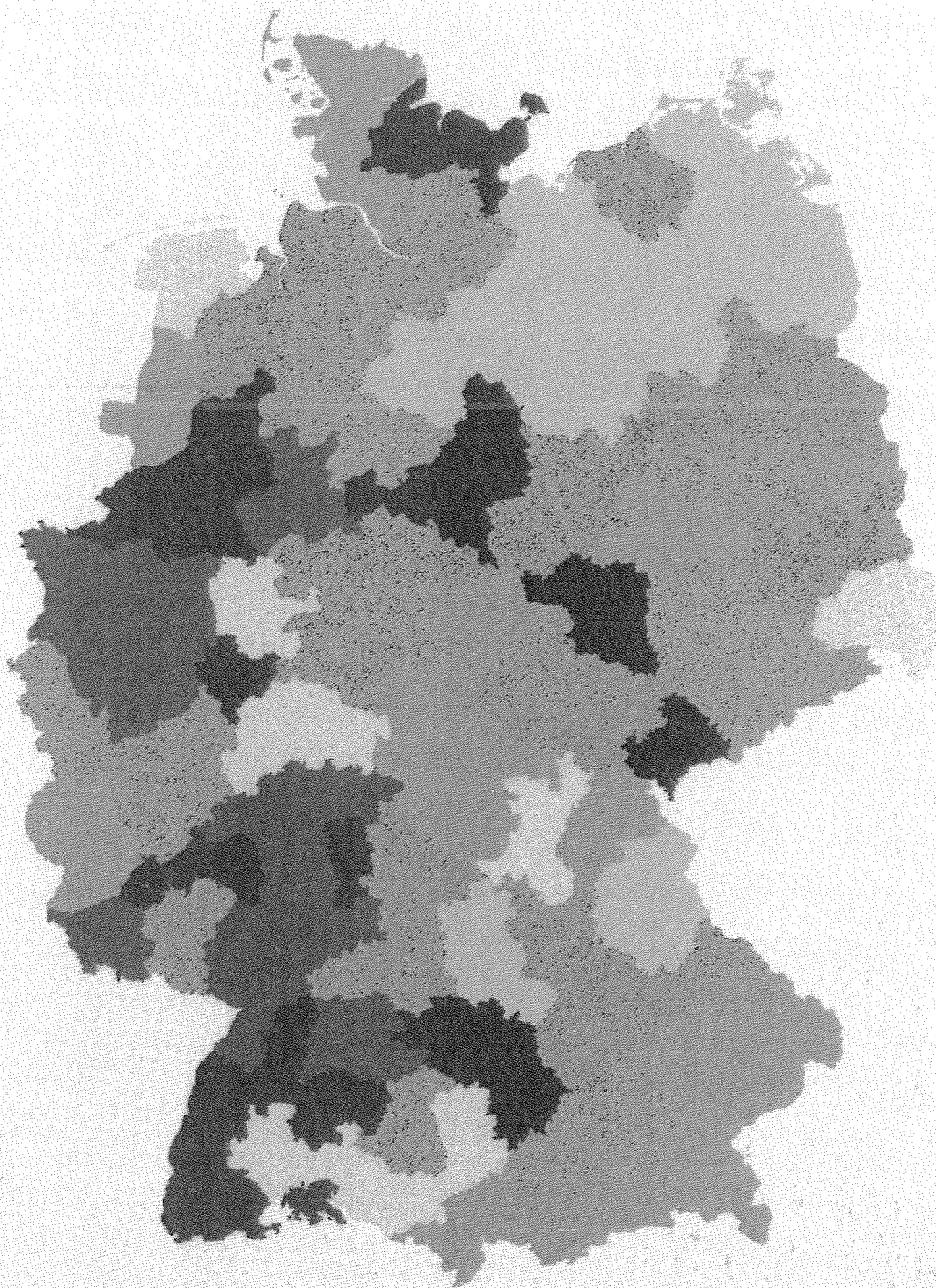


Gerd Grözingen, Wenzel Matiaske (Hrsg.)

Deutschland regional

Sozialwissenschaftliche Daten im Forschungsverbund



Veröffentlichung mit freundlicher Genehmigung durch
den Rainer Hampp Verlag

Rainer Hampp Verlag

Das Sozio-oekonomische Panel (SOEP) und die Möglichkeiten regionalbezogener Analysen

C. Katharina Spieß*

1 Allgemeines zum SOEP

Das Sozio-oekonomische Panel (SOEP) ist eine repräsentative Wiederholungsbefragung privater Haushalte in Deutschland. Für das SOEP wird im jährlichen Rhythmus seit 1984 bei denselben Personen und Haushalten in der Bundesrepublik eine standardisierte Erhebung durchgeführt. Im Juni 1990 wurde die Studie auf das Gebiet der ehemaligen DDR ausgeweitet. Zur adäquaten Erfassung des gesellschaftlichen Wandels wurde in den Jahren 1994/95 die sogenannte „Zuwanderer-Stichprobe“ eingeführt. In den Jahren 1998 und 2000 wurde die Stichprobengröße verdoppelt. Seit 2002 sind die Bezieher hoher Einkommen überrepräsentiert. Mit Hilfe des SOEP können die politischen und gesellschaftlichen Veränderungen beobachtet und analysiert werden, denn der Datensatz gibt sowohl Auskunft über objektive Lebensbedingungen als auch über subjektiv wahrgenommene Lebensqualität, über den Wandel in verschiedenen Lebensbereichen und über die Abhängigkeiten, die zwischen verschiedenen Lebensbereichen und deren Veränderungen existieren. Die SOEP-Daten werden jährlich von über 400 Nutzern analysiert, daraus ca. 150 Nutzer im Ausland. Auf Basis der SOEP-Daten entstehen laufend Publikationen in den besten Zeitschriften der Sozial- und Wirtschaftswissenschaften.¹

Die Stärken des SOEP bestehen insbesondere in seinen besonderen Analysemöglichkeiten durch das Längsschnittdesign (Panelcharakter). Darüber hinaus ist es ein großer Vorteil des SOEP, dass eine Befragung aller erwachsenen Haushaltsmitglieder erfolgt, die Möglichkeit innerdeutscher Vergleiche gegeben ist und durch eine überproportionale Ausländerstichprobe eigene Analysen für die ausländische Bevölkerung in Deutschland möglich sind. Gegenwärtig ist das SOEP die größte Wiederholungsbefragung bei Ausländern in der Bundesrepublik Deutschland; die Spezialstichprobe umfasst Haushalte mit einem Haushaltsvorstand türkischer, spanischer, italienischer, griechischer oder ehemals jugoslawischer Nationalität. Ferner ist die Erhebung von Zuwanderung gegenwärtig die einzige metho-

* Ich danke Gert G. Wagner, DIW Berlin für hilfreiche Anmerkungen zum Manuskript dieses Beitrags.

¹ Vgl. das Informationssystem über Publikationen auf Basis des SOEP,
<http://www.diw.de/deutsch/sop/soeppub/soeplit/>

disch zuverlässige Stichprobe von Zuwanderern, d.h. auch von Aussiedlern, die von 1984 bis 1995 nach Westdeutschland gekommen sind.

Das SOEP zeichnet sich durch eine hohe Stabilität aus. 1984 beteiligten sich im SOEP-West 5.921 Haushalte mit 12.290 erfolgreich befragten Personen an der Erhebung; nach 19 Wellen im Jahre 2002 sind es in den Originalstichproben noch 3.889 Haushalte mit 7.175 Personen. Im SOEP-Ost wurden 1990 2.179 Haushalte mit 4.453 Personen befragt; 2002 gaben 3.466 Personen in 1.818 Haushalten Auskunft über ihre Lebenssituation. Die Zuwanderer-Stichprobe der Jahre 1994/95 behielt mit 402 Haushalten und 780 Personen im Jahre 2002 einen recht stabilen Umfang. Von den 1.910 Personen in 1.056 Haushalten einer Ergänzungsstichprobe aus dem Jahre 1998 konnten im Jahre 2002 1.373 Personen in 773 Haushalten erneut befragt werden. 2000 wurde eine Innovationsstichprobe gezogen, um auf Basis einer großen Fallzahl bessere Analysen kleiner Teilgruppen der Bevölkerung zu ermöglichen. Für das Jahr 2002 stehen aus dieser Stichprobe Informationen zu weiteren 8.427 Befragungspersonen in 4.586 Haushalten zur Verfügung. Im Jahre 2002 konnte schließlich eine Zufalls-Stichprobe für Haushalte von Hocheinkommensbezieherern realisiert werden. Sie umfasst 1.224 Haushalte mit 2.671 Befragungspersonen. Im Jahr 2006, dem wahrscheinlichen Startjahr einer Stichprobenauffrischung, werden ca. 21.200 Personen in 11.300 Haushalten befragt werden.

Mit denen im SOEP erhobenen Informationen wird ein breites Themenspektrum abgedeckt. Das SOEP liefert kontinuierlich Informationen u.a. über Haushaltszusammensetzung, Wohnsituation, Erwerbs- und Familienbiographien, Erwerbsbeteiligung und berufliche Mobilität, Einkommensverläufe, Gesundheit, Lebenszufriedenheit und gesellschaftliche Partizipation. Es werden objektive und subjektive Indikatoren erfasst. In jährlich wechselnden Schwerpunktthemen werden darüber hinaus Informationen bereitgestellt, z.B. über Familie und soziale Dienste, Zeitverwendung, Weiterbildung und Qualifikation sowie die soziale Sicherung.

Die Durchführung und Entwicklung der Untersuchung liegt bei der Serviceeinrichtung "Längsschnittstudie Sozio-oekonomisches Panel" der Leibnitz-Gemeinschaft (WGL). Das SOEP ist als eine selbständige Abteilung in das DIW Berlin integriert. Die Serviceeinrichtung SOEP gibt die Daten an die interessierte Fachöffentlichkeit weiter und erstellt eigene Analysen. Die Feldarbeit führt Infratest Sozialforschung, München durch. Das SOEP kann prinzipiell von allen Wissenschaftlerinnen und Wissenschaftlern genutzt werden. Der anonymisierte Mikro-Datensatz des SOEP wird Universitäten und anderen Forschungseinrichtungen für Forschung und Lehre gegen eine geringe Datenweitergabegebühr von 25 Euro zur Verfügung gestellt. Die SOEP-Daten werden als Rohdaten sowie in SPSS-, SAS-, STATA- und TDA-Format mit ausführlichen Arbeitsunterlagen auf CD-ROM weitergegeben. Eine Vielzahl weiterer Informationen zum SOEP sind der Homepage des SOEP zu entnehmen, vgl. <http://www.diw.de/deutsch/soep> oder auch den folgenden Beiträgen Wagner et al. (1993), Schupp und Wagner (2002) sowie SOEP Group (2001).

2 Regionalbezogene Analysen mit dem SOEP

Das SOEP bietet vielfältige Möglichkeiten regionalbezogener Analysen. Über die regionale Zuordnung des Wohnortes der SOEP-Haushalte bzw. der Personen in diesen Haushalten ist es möglich eine Vielzahl von regionalen Indikatoren auf der Ebene der Bundesländer, der Raumordnungsregionen, der Kreise und der Postleitzahlen dem SOEP zuzuspielen, wobei wegen der datenschutzrechtlichen Sensibilität dieser Regionalinformationen bestimmte Sicherheitsvorkehrungen einzuhalten sind.

2.1 Regionalinformationen im Standarddatensatz

Dem Standarddatensatz des SOEP ist das Bundesland zu entnehmen, in dem ein SOEP-Haushalt wohnt. Außerdem ist die Gemeindengröße des Wohnortes bekannt. Im Standarddatensatz des SOEP sind der Gemeindetyp nach Bousted und BIK sowie die politische Gemeindegrößeklasse enthalten. In einigen Erhebungsjahren wird außerdem die Entfernung zur nächsten Großstadt oder zu anderen Infrastruktureinrichtungen (Geschäfte, Banken, Ärzte, Schulen, Sportanlagen, Verkehrsmittel, Gaststätten etc.) und die Siedlungsform (d.h. die Einordnung als Großstadt, Industriegebiet oder ländliches Gebiet) erhoben. Außerdem wird in einigen Wellen des SOEP die Wohngegend darüber beschrieben, inwiefern beispielsweise eine Beeinträchtigung durch eine Lärmbelästigung, Luftverschmutzung oder den Mangel an Grünflächen stattfindet. Es wird erfragt, ob die Wohngegend sich durch einen hohen Ausländeranteil oder durch einen hohen Anteil von Landleuten charakterisieren lässt.²

2.2 Die Ebene der Raumordnungsregionen

Die Raumordnungsregion, in der ein SOEP-Haushalt lebt, muss aus datenschutzrechtlichen Gründen gesondert angefordert werden. Für Analysen mit diesen Daten muss der SOEP-Gruppe ein spezielles Datenschutzkonzept vorgelegt werden. Ist dies der Fall, werden den SOEP-Nutzern die Raumordnungsregionen gesondert bereit gestellt.

Wie der Tabelle 1 zu entnehmen ist, werden über die Wohnsitze der SOEP-Haushalte alle Raumordnungsregionen repräsentiert. Im Mittel sind seit der Erweiterung der SOEP-Stichprobe im Jahr 2000 in einer Raumordnungsregion 200 Haushalte enthalten, wobei es Raumordnungsregionen gibt, in denen über 500 SOEP-Haushalte leben und andere Raumordnungsregionen, die nur mit 30 SOEP-Haushalten repräsentiert sind.

² In den ersten Wellen neuer Stichproben wird das Wohnumfeld der Haushalte über eine Quartiersbeschreibung der Interviewerinnen und Interviewer erfasst.

Tabelle 1: Kennziffern zur Analyse mit Raumordnungsregionen (ROR) im SOEP (1985-2002)

	19 85	19 86	19 87	19 88	19 89	19 90	19 91	19 92	19 93	19 94	19 95	19 96	19 97	19 98	19 99	20 00	20 01	20 02
Anzahl der im SOEP repräsentier- ten ROR's	75	75	75	75	75	97	97	97	97	97	97	97	97	97	97	97	97	97
Mittlere Anzahl der HH pro ROR	123	117	116	111	108	127	121	123	122	118	126	118	115	126	120	209	190	207
Maximale Anzahl der HH pro ROR	281	266	267	256	243	454	411	427	415	399	424	327	312	334	318	530	494	521
Minimale Anzahl der HH pro ROR	7	7	7	7	7	7	6	6	7	7	14	15	11	13	14	34	26	31
Basis: nur Haushalte mit realisiertem Haushaltsinterview																		
Quelle: SOEP 1994-2002, eigene Berechnungen																		

Bei Längsschnittbetrachtungen muss miteinbezogen werden, dass sich die Raumordnungsregionen der Jahre 1985 bis 1995 auf einen anderen Gebietsstand beziehen als die Raumordnungsregionen seit 1996, gleichwohl die Zahl der Raumordnungsregionen seit 1990 mit insgesamt 97 Regionen konstant geblieben ist. Längsschnittanalysen über diese beiden Zeiträume hinweg sind also nur bedingt möglich.³

Auf der Ebene der Raumordnungsregionen können Informationen der laufenden Raumbeobachtung des Bundesamtes für Bauwesen und Raumordnung (BBR)⁴ in Bonn zugespielt werden. Die Regionalindikatoren liegen zum einen in schriftlicher Form vor, vgl. dazu die einschlägige Reihe des BBR.⁵ Zum anderen stellt das BBR seit der zweiten Hälfte der 90er Jahre die Regionalindikatoren auch in elektronischer Form der breiten Fachöffentlichkeit zur Verfügung.

³ Für nähere Informationen zur Neuabgrenzung der Raumordnungsregionen, vgl. z.B. Böltken (1996).

⁴ Vgl. <http://www.bbr.bund.de>.

⁵ Vgl. die diversen Bände *Bundesamt für Bauwesen und Raumordnung (Hg.): Aktuelle Daten zur Entwicklung der Städte, Kreise und Gemeinden, Berichte des BBR*, Bonn.

2.3 Die Ebene der Kreise und Postleitzahlen

Für SOEP-Nutzer, welche die Information benötigen, in welchem Kreis ein SOEP-Haushalt lebt, ist es möglich am DIW Berlin entsprechende Analysen durchzuführen. Ab dem Jahr 2005 wird es außerdem möglich sein, kreisbezogene Analysen mit dem SOEP vom Arbeitsplatz der SOEP-Nutzer, d.h. auch außerhalb des DIW Berlin, durchzuführen (vgl. dazu Kapitel 3).

Wie Tabelle 2 für die SOEP-Wellen ab 1985 darstellt, werden ab dem Jahr 2000 mit über 430 Kreisen fast alle deutschen Kreise im SOEP repräsentiert – gegenwärtig existieren 440 Landkreise und Stadtkreise. Durchschnittlich sind in einem Kreis knapp 60 Haushalte vertreten. Mit fast 300 Haushalten ist Berlin als kreisfreie Stadt der Kreis mit den meisten SOEP-Haushalten.

Tabelle 2: Kennziffern zur Analyse mit Kreiskennziffern im SOEP

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Anzahl der im SOEP repräsentierten Kreise	278	288	292	292	292	455	458	465	470	477	495	441	422	425	424	435	436	437
Mittlere Anzahl der HH pro Kreis	45	43	42	40	39	36	35	35	34	33	34	35	34	38	36	60	55	59
Maximale Anzahl der HH pro Kreis	204	200	198	189	187	182	179	177	174	169	196	191	173	199	187	298	281	293
Anmerkung: Die minimale Anzahl der Kreise kann aus Gründen des Datenschutzes nicht ausgewiesen werden.																		
Basis: nur Haushalte mit realisiertem Haushaltsinterview																		
Quelle: SOEP 1994-2002, eigene Berechnungen																		

Für Längsschnittanalysen, welche Kreisinformationen nutzen, müssen die Einschränkungen bzw. Probleme beachtet werden, die mit den Gebietsreformen (Kreisgebietsreformen und Gemeindegebietsreformen) Ostdeutschlands verbunden sind. Da sich durch diese Gebietsreformen in den Jahren 1993 bis 1996 die Gebietsstände der Kreise und Gemeinden verändert haben, ist seit dem Beginn der Gebietsreformen eine kontinuierliche Betrachtung der Kreise über die Zeit kaum mehr möglich. Für Längsschnittanalysen kommt erschwerend hinzu, dass die Gebietsreformen (Kreisgebietsreformen) in den einzelnen ostdeutschen Bundes-

ländern zu unterschiedlichen Zeitpunkten begonnen und insbesondere zu unterschiedlichen Zeitpunkten abgeschlossen wurden.⁶

Auf der Ebene der Kreiskennziffern können den SOEP-Daten unterschiedliche Indikatoren zugespielt werden. Das Statistische Bundesamt in Zusammenarbeit mit den Statistischen Landesämtern stellt mit den CD-Roms „Statistik regional“ und „Statistik lokal“⁷ eine Vielzahl von Indikatoren auf Kreisebene zur Verfügung.

Unter sehr besonderen Datenschutzmaßnahmen können auch Informationen auf der Ebene der fünfstelligen Postleitzahlen benutzt werden. Da es sich dabei um datenschutzrechtlich sehr sensible Daten handelt, dürfen auch diese nicht nach außen gegeben werden und sie stehen auch nicht über SOEP-Remote (vgl. Kapitel 3) zur Verfügung. SOEP-bezogene Analysen auf der Ebene der Postleitzahlen können bis auf weiteres nur am DIW Berlin an einem speziell dafür eingerichteten Arbeitsplatz durchgeführt werden.

Die Postleitzahlen sind im SOEP ab dem Jahr 1993 verfügbar. Auf dieser Ebene können Informationen zugespielt werden. Allerdings handelt es sich bei den Postleitzahlen um keine amtliche Gebietsabgrenzung, sondern die Postleitzahlbezirke sind regionale Einheiten der Deutschen Post AG, die vorrangig zur Organisation der Briefzustellung gebildet werden. Insbesondere kommerzielle Datenanbieter bieten auf dieser Ebene Daten zum Beispiel zur Kaufkraft, der Bebauung oder der Firmenstruktur an (vgl. dazu die Beiträge von Rolf Küppers, microm und Bernd Lochschmidt, GfK in diesem Band).

3 Bisherige regionalbezogene Arbeiten, praktische Hinweise und Fazit

Untersuchungen, welche die leicht zugänglichen regionalbezogenen Informationen des SOEP nutzen, liegen zu unterschiedlichen Themenschwerpunkten vor, vgl. dazu das Informationssystem über Publikationen auf Basis des SOEP. Bisherige Analysen, welche die kreisbezogenen Informationen des SOEP nutzen, umfassen zum Beispiel die Arbeiten von Rehdanz und Maddison (2004), die mit dem SOEP Klimadaten auf Kreisebene verknüpfen, und von Baumgartner (2004), der dem SOEP Informationen über die Größe von Schulklassen zuspielt. Die Arbeiten von Büchel und Spieß (2002), Hank (2003a und b), Hank und Kreyenfeld (2003), Hank et al. (2004), Kreyenfeld und Hank (2000), Spieß (1998), Spieß und Büchel (2003) sowie Wrohlich (2004) verbinden mit dem SOEP Informationen der Kinder- und Jugendhilfestatistik.

Arbeiten, welche die Information über den Postleitzahlbezirk des SOEP-Haushaltes nutzen, existieren von Drever (2004), die den SOEP-Daten Informationen über die regionale Kaufkraft und den Migrantenanteil eines Postleitzahlbezirks zuspielt. Außerdem können die Postleitzahlen zur Abgrenzung von Arbeitsamtbe-

⁶ Zu den Gebietsreformen und ihren Folgen für Längsschnittsuntersuchungen, vgl. z.B. Blach und Jonetzko (1999).

⁷ Vgl. <http://www.statistik-portal.de/Statistik-Portal/publ.asp>.

zirken verwendet werden, vgl. für eine solche SOEP-basierte Analyse die Arbeit von Silke Anger in diesem Band.

Doch nicht nur für die Verknüpfung mit Indikatoren, die sich auf unterschiedliche regionale Einheiten beziehen, sind die Informationen über unterschiedliche regionale Einheiten der Wohnorte von SOEP-Haushalten von Interesse. Auch zur Berechnung von Entfernungen ist die Kenntnis der regionalen Zugehörigkeit der SOEP-Haushalte bedeutsam; vergleiche dazu zum Beispiel die Arbeit von Johannes Schwarze in diesem Band. In dieser Untersuchung werden die Entfernungen zwischen dem Wohnort von Eltern und Kindern im SOEP berechnet. Auch für die Analyse von Wohnortwechseln sind die Informationen zur kleinräumigen Zugehörigkeit von SOEP-Haushalten von Interesse, vgl. z.B. Hunt (2004). Darüber hinaus ist die regionale Zugehörigkeit eines SOEP-Haushalts auch für ökonometrische Analysen von Bedeutung, da darüber für kleinräumige regionale Besonderheiten, z.B. auf der Ebene der Kreise oder der Postleitzahlen, kontrolliert werden kann.

Als Fazit lässt sich festhalten, dass das SOEP vielfältige Möglichkeiten regionalbezogener Analysen bietet. Diese werden künftig durch verbesserte Zugangsmöglichkeiten zu datenschutzrechtlich sensiblen Daten für eine noch breitere wissenschaftliche Nutzergemeinschaft von Interesse sein. Diese verbesserten Zugangsmöglichkeiten werden es Nutzern ab dem Jahr 2005 ermöglichen, über SOEP-Remote von ihrem Arbeitsplatz kreisbezogene Analysen durchzuführen. Über SOEP-Remote können automatisiert und via E-Mail externe Datenjobs, d.h. von Nutzern außerhalb des DIW Berlin, über die datenschutzrechtlich sensiblen Daten laufen. Dabei haben externe Nutzer physisch keinen Zugang zu den sensiblen Daten. Ihnen wird lediglich das Ergebnis ihrer Analysen automatisiert und via E-Mail zur Verfügung gestellt. Für die neusten Entwicklungen, die SOEP-Remote betreffen, vgl. die Homepage der SOEP-Gruppe.

Literatur

- Baumgartner, Hans J. (2004): Are There Any Class Size Effects On Early Career Earnings in West Germany ?, DIW Discussion Papers 417, Berlin.
- Blach, Antonia und Jacek Jonetzko (1999): Die Gebietsreform der neuen Länder: Folgen für die Laufende Raumbbeobachtung des BBR, Arbeitspapier 5/1999 des Bundesamtes für Bauwesen und Raumordnung, Bonn.
- Böltken, Ferdinand et. al. (1996): Neuabgrenzung von Raumordnungsregionen nach den Gebietsreformen in den neuen Bundesländern, Arbeitspapier 5/1996 der Bundesforschungsanstalt für Landeskunde und Raumordnung, Bonn.
- Büchel, Felix und C. Katharina Spieß (2002): Müttererwerbstätigkeit und Kindertageseinrichtungen – neue Ergebnisse zu einem bekannten Zusammenhang, in: Vierteljahrshefte zur Wirtschaftsforschung (71: 96-114).

- Drever, Anita I. (2004): Separate Spaces, Separate Outcomes? Neighbourhood Impacts on Minorities in Germany, in: *Journal of Urban Studies* (41: 1423-1439).
- Hank, K. (2003a): The Differential Influence of Women's Residential District on the Risk of Entering First Marriage and Motherhood in Western Germany, in: *Population and Environment* (25: 3-21).
- Hank, K. (2003b): Eine Mehrebenenanalyse regionaler Einflüsse auf die Familiengründung westdeutscher Frauen in den Jahren 1984 bis 1999, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (55: 79-98).
- Hank, K. und M. Kreyenfeld (2003): A Multilevel Analysis of Child Care and Women's Fertility Decisions in Western Germany, in: *Journal of Marriage and Family*, (65: 584-596).
- Hank, K., M. Kreyenfeld und C. K. Spieß (2004): Kinderbetreuung und Fertilität in Deutschland, in: *Zeitschrift für Soziologie* (33: 228-244).
- Hunt, Jenny (2004): Are Migrants More Skilled than Non-Migrants? Repeat, Return and Same-Employer Migrants, in: *Canadian Journal of Economics* (37: 830-849).
- Kreyenfeld, M. und K. Hank (2000): Does the availability of child care influence the employment of mothers? Findings from western Germany, in: *Population Research and Policy Review* (19: 317-337).
- Rehdanz, K. and Maddison, D. (2004): The Amenity Value of Climate to German Households, FNU-39, Centre for Marine and Climate Research, Hamburg University, Hamburg. Auch erschienen als Nota di Lavoro 57.2004, Milano: Fondazione Eni Enrico Mattei, Milano und als DIW Discussion Paper 414, German Institute for Economic Research, Berlin.
- Schupp, Jürgen und Gert G. Wagner (2002): Maintenance of and Innovation in Long-term Panel Studies. The Case of the German Socio-Economic Panel (GSOEP), in: *Allgemeines Statistisches Archiv* (86: 163-175).
- SOEP Group (2001): The German Socio-Economic Panel (GSOEP) after more than 15 years – Overview, in: *Vierteljahrshefte zur Wirtschaftsforschung* (70: 7-14).
- Spieß, C. Katharina und Felix Büchel (2003): Effekte der regionalen Kindergarteninfrastruktur auf das Arbeitsangebot von Müttern, in: *Soziale Sicherung und Arbeitsmarkt* (hrsg. von W. Schmähl), Reihe „Schriften des Vereins für Socialpolitik“, Duncker & Humblot: Berlin 2003, S. 95-126.
- Wagner, Gert G., Burkhauser, Richard und Friederike Behringer (1993): The English Language Public Use File of the German Socio-Economic Panel, in: *The Journal of Human Resources* (28: 429-433).
- Wrohlich, Katharina (2004): Child Care Costs and Mothers' Labor Supply: An Empirical Analysis for Germany, DIW Discussion Papers 412, Berlin.