

Knies, Gundi; Spieß, C. Katharina

**Research Report**

## Regional Data in the German Socio-Economic Panel Study (SOEP)

DIW Data Documentation, No. 17

**Provided in Cooperation with:**

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* Knies, Gundi; Spieß, C. Katharina (2007) : Regional Data in the German Socio-Economic Panel Study (SOEP), DIW Data Documentation, No. 17, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/129226>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



**DIW** Berlin

Deutsches Institut  
für Wirtschaftsforschung

## Data Documentation 17



Gundi Knies  
C. Katharina Spiess

**Regional Data in the German Socio-Economic  
Panel Study (SOEP)**

## IMPRESSUM

© DIW Berlin, 2007

DIW Berlin  
Deutsches Institut für Wirtschaftsforschung  
Königin-Luise-Str. 5  
14195 Berlin  
Tel. +49 (30) 897 89-0  
Fax +49 (30) 897 89-200  
[www.diw.de](http://www.diw.de)

ISSN 1861-1532

All rights reserved.  
Reproduction and distribution  
in any form, also in parts,  
requires the express written  
permission of DIW Berlin.



## **Data Documentation 17**

Gundi Knies \*

C. Katharina Spiess \*\*

### **Regional Data in the German Socio-Economic Panel Study (SOEP)**

Berlin, January 2007

\* DIW Berlin, SOEP, and CMPO, University of Bristol

\*\* DIW Berlin, SOEP, and FU Berlin, [kspiess@diw.de](mailto:kspiess@diw.de)



## Contents

<b>1</b>	<b>Introduction to regional information in SOEP .....</b>	<b>1</b>
1.1	Official geographical units in SOEP .....	2
<b>2</b>	<b>SOEP data at the regional policy region level.....</b>	<b>5</b>
2.1	Implications of county restructuring in East Germany 1993-1996 on ROR .....	5
2.2	Implications of ROR readjustments in 1996.....	7
<b>3</b>	<b>SOEP data at the county level .....</b>	<b>9</b>
3.1	Information on recoded county codes.....	9
3.2	Additional information at the county level .....	11
<b>4</b>	<b>SOEP data at the zip-code level .....</b>	<b>14</b>
4.1	Geographical reference of zip codes.....	14
4.2	SOEP sample distributions by zip-code areas .....	15
4.3	Potential for Analyses of SOEP data at the zip-code level.....	19
	Neighbourhood indicators.....	19
	Research questions.....	21
	A cautionary note.....	21
<b>5</b>	<b>Further information at the neighbourhood level.....</b>	<b>23</b>
5.1	Indicators of the built environment.....	23
5.2	Indicators of the local infrastructure and the natural environment .....	24
5.3	Indicators of the social environment.....	26
	<b>References .....</b>	<b>28</b>
	<b>Appendix .....</b>	<b>29</b>

## **Contents of Tables and Figures**

Table 1	Representation of regional policy regions (ROR) in GSOEP. Number of observations, 1984-2005. ....	6
Table 2	Representation of counties (Kreise and kreisfreie Städte) in GSOEP. Number of observations, 1985-2005 .....	10
Table 3	Selected indicators at county level in SOEP, 2002 .....	12
Table 4	Representation of unique zip-code areas in SOEP, 1993, 1998 and 2004 .....	16
Table 5	Number of zip-codes represented in SOEP differentiated by zip-code region, 1993-2004.....	17
Table 6	Number of households in SOEP differentiated by zip-code regions, 1993-2004 .....	18
Table 7	Number of individuals in SOEP differentiated by zip-code regions, 1993-2004 .....	18
Table 8	Number of households in SOEP who still lived in the same zip-code area in the subsequent wave and five years later .....	20
Table 9.1	Number of SOEP households differentiated by how long it takes to walk to local basic amenities. 1986, 1994, 1999 and 2004 .....	24
Table 9.2	Number of SOEP households differentiated by how long it takes to walk to local facilities for particular age groups, 1986, 1994, 1999 and 2004 .....	25
Table 9.3	Number of SOEP households differentiated by how long it takes to walk to local facilities for leisure activities, 1986, 1994, 1999 and 2004 .....	25
Table 9.4	Number of SOEP households differentiated by how much they are bothered by noise pollution, pollution and lack of access to green space, 1986, 1994, 1999 and 2004 .....	26
Table 10	Presence of foreigners in residential areas 1986, 1994, 1999 and 2004 .....	26
Box 1	Zip codes – the meaning of the digits.....	15
App. 1	Map of Federal States of Germany.....	29
App. 2	Map of Regional Policy Regions (ROR) and Counties in Germany, 2004.....	30
App. 3	Map of zip-code regions in Germany (first digit identifies zip-code master region) .....	31
App. 4	Map of counties in Germany, differentiated by number of SOEP households observed in 2004 .....	32
App. 5	Map of counties in Germany, differentiated by number of SOEP individuals observed in 2004 .....	33
App. 6	Map of counties in Germany, differentiated by number of SOEP children observed in 2004 .....	34

## 1 Introduction to regional information in SOEP

Throughout the 1990s, social scientists showed a growing interest in geography as an explanatory factor for social inequalities. In this data documentation, we first describe the geographical data currently available for use with the German Socio Economic Panel Study (SOEP) and discuss how these data can be operationalised by researchers using the SOEP data.

The SOEP contains a great deal of geographically referenced indicators and regional information. First, SOEP contains variables that indicate the area in which survey respondents live at the time of the (household) interview. The values of these variables refer to *official geographical units* used either by German administrative bodies or by the postal service (Deutsche Post AG) to deliver mail to households in Germany. On the basis of these unique identifiers it is possible, in principle, to match SOEP data with official, scientific or commercial macro-data at these levels.

Second, the SOEP collects a great deal of *information on the local environment* in which its respondents live. These indicators are collected from the respondents themselves and do not refer to geographical entities delineated by German administrative bodies. The indicators do not refer to any systematically delineated area but simply to whichever area the respondents consider their own ‘residential area’, ‘neighbourhood’, or ‘place of residence’.

Third, SOEP contains *select macro-indicators at the scale of select geographical units* available in the context of the panel study. The provision of these macro-indicators is the result of efforts by the SOEP group to match SOEP with official data at different geographical scales. This has been undertaken to ease future matching with macro-indicators at these levels.

The use of the first type of indicators with SOEP is subject to data protection restrictions. Since geographical unit identifiers refer to officially defined areas in Germany, indicators at these smaller geographical scales are not distributed with the standard SOEP CD-ROM. They are provided only under a separate contract subject to binding regulations.

In the remainder of this section, we introduce researchers using SOEP to the official geographical units represented in SOEP, and provide a quick guide to the regulations that apply to the use of these geographically referenced indicators for analyses with SOEP.



## 1.1 Official geographical units in SOEP

The official geographical units assigned in SOEP are federal states (*Bundesländer*), regional policy regions (*Raumordnungsregionen*, ROR), administrative regions (*Regierungsbezirke*), counties (*Kreise* and *kreisfreie Städte*), municipalities (*Gemeinden*) and zip-code areas (*Postleitzahlengebiete*).

The broadest level at which SOEP respondents can be differentiated and spatial indicators can be matched is that of the German **federal states**, or *Länder*. There are 16 federal states in Germany (see map in Appendix 1). Information about each of the federal states in which SOEP respondents live is included in the standard data set. This indicator can also be used to distinguish people living in the East from people living in the West. German federal state identification numbers in principle correspond to NUTS I identification numbers at the European level.

The SOEP data on the federal states should not be taken as representative of the entire population within each area. SOEP is representative at the national level but may not necessarily be representative at the level of each individual state (or at the geographically smaller regional level). The only states that may be analysed individually are the large states with small confidence bands: North Rhine-Westphalia, Bavaria, and in some cases Baden-Wuerttemberg. However, due to small case numbers, it may not be possible to draw broader, statistically significant conclusions from detailed structural analyses of individual cells. To avoid this problem, data on states with similar characteristics or neighbouring states may be pooled.<sup>1</sup> Owing to a substantive sample extension in 2000, new possibilities have been opened up for SOEP analyses. The information on the federal states is a standard variable in SOEP and no further data protection regulations apply.

**Regional policy regions** (*Raumordnungsregionen*, ROR), are spatial units defined by the Federal Office for Building and Regional Planning (*Bundesamt für Bauwesen und Raumordnung*, BBR<sup>2</sup>) to differentiate areas in Germany based on their economic interlinkages. There are 97 different regional policy regions in Germany (see map in Appendix 2). Indicators at this geographical level represent the core element of the SOEP geocode data module of SOEP. For SOEP users in Europe, access to this module containing regional information is

---

<sup>1</sup> See Frick and Goebel (2005) for an example.

<sup>2</sup> <http://www.bbr.bund.de>

conditional on submitting a special data protection plan and signing a data distribution contract. SOEP users living overseas can be provided access to the regional SOEP geocode data via SOEP-remote.<sup>3</sup>

The third geographical unit at which analysis and matching of SOEP may be undertaken is the level of **administrative regions** (*Regierungsbezirke*). The number of administrative regions in Germany amounted to 22 in the year 2005. In some states, NUTS II identifiers at the European level correspond to these administrative regions, while in others, the NUTS II identifiers correspond to other regional units. Information on the administrative region in which SOEP respondents live has been provided since 1985. These data can only be provided to users on the DIW Berlin premises, for example in the framework of a research visit.

SOEP also provides data at yet another, geographically smaller regional scale: the **county level** (*Kreise* and *kreisfreie Städte*). Counties are comparable with but not necessarily identical to NUTS III at the European level. Data at the county level can be provided on the premises of DIW Berlin as well as via SOEP-remote. Select county-level indicators are provided in the SOEP county-level module. This level may also be matched to researchers' own indicators.

Since 2000, SOEP contains select information at the **level of the municipality** (*Gemeinde*) in which the respondents of the survey live. On the basis of the identification number of the municipalities it is also possible to match SOEP with other data at the municipal level.<sup>4</sup> These data can only be analysed on the premises of DIW Berlin.

The smallest regional unit provided in SOEP is the **zip-code level**. Researchers can work with SOEP zip-code data only on the premises of DIW Berlin.

In addition to the regional identification numbers and select indicators at these geographical levels, SOEP provides a great deal of further information on the regional contexts in which SOEP respondents live. This includes BOUSTEDT or BIK indicators and, among other things, distance to public facilities in the area (the latter will be described below). All these indicators are provided in the standard SOEP data. However, users who want to work with variables referring to BOUSTEDT or BIK (such as *city size*) indicators need to sign an addi-

---

<sup>3</sup> SOEP-remote is a special databank which allows users outside the premises of DIW Berlin to obtain results of county level data without having physical access to the sensitive data at this regional scale (see Goebel 2005).

<sup>4</sup> For instance, the SOEP team uses an indicator of housing prices at this level to impute missing data on wealth. This is documented in Frick and Grabka (2007).

tional data use contract. The SOEP team then provides a special password allowing extraction of this information from the standard SOEP CD-ROM.

In the following sections we discuss the SOEP data in more detail at the level of regional policy regions (Section 2), counties (Section 3) and zip-code areas (Section 4). Section 5 provides a survey of the information available about the local environments in which SOEP respondents live. In each of these sections, we present raw frequencies of the regional units represented in SOEP (in both the cross-sectional and longitudinal perspective), and point out potentials and limitations for socio-scientific analyses.

## 2 SOEP data at the regional policy region level

Information at the level of Regional Policy Regions (*Raumordnungsregionen*, ROR) is available for all waves of SOEP from 1985 onwards. All 97 ROR in Germany are represented in the SOEP. In 2005, SOEP observed an average of 184 households per region, with 470 being the maximum and 27 the minimum (see Table 1).

The SOEP geocode data module contains the official identification numbers of the regional policy regions (ROR) as defined by Federal Office for Building and Regional Planning (BBR, previously BfLR). In addition, it contains a number of select macro-indicators that have been provided by BBR. SOEP matched with regional indicators provided by BBR is available for the years 1984 to 1994 and for West Germany only. Researchers wishing to analyse more recent waves of SOEP (or who want to draw on different macro-indicators) may match the survey with different or more recent BBR indicators. Since 1995 indicators can be obtained directly from BBR (see below).

### 2.1 Implications of county restructuring in East Germany 1993-1996 on ROR

The SOEP team determines respondents' regional policy regions based on the counties in which they live. Longitudinal analyses of SOEP geocode data are subject to limitations due to official changes in the boundaries of counties and municipalities in East Germany between 1993 and 1996 (undertaken as part of the “*Gebietsreform der neuen Länder*” reform program redefining regional boundaries). Overall, the number of counties in East Germany decreased from 215 in 1991 to 111 in 1996.

Table 1

Representation of regional policy regions (ROR) in GSOEP. Number of observations, 1984-2005.

<b>1984-1994</b>	<b>1985</b>	<b>1986</b>	<b>1987</b>	<b>1988</b>	<b>1989</b>	<b>1990</b>	<b>1991</b>	<b>1992</b>	<b>1993</b>	<b>1994</b>	<b>1995</b>
<b>Number of ROR</b>	75	75	75	75	75	97	97	97	97	97	97
<b>Number of households in ROR</b>											
<i>mean</i>	122.8	117.05	115.86	111.09	108.35	126.96	121.31	122.92	121.74	113.5	118.4
<i>maximum</i>	281	266	267	256	243	454	411	428	416	343	340
<i>minimum</i>	7	7	7	7	7	7	6	6	7	7	14
<b>1995-2005</b>	<b>1996</b>	<b>1997</b>	<b>1998</b>	<b>1999</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	
<b>Number of ROR</b>	97	97	97	97	97	97	97	97	97	97	
<b>Number of households in ROR</b>											
<i>mean</i>	117.34	114.47	125.96	120.07	208.97	190.36	207.38	198.05	191.02	184.39	
<i>maximum</i>	327	312	334	318	530	494	521	508	500	470	
<i>minimum</i>	15	11	13	14	34	26	31	29	30	27	

Source: SOEP 21. Authors' calculations.

Notes: Only households/persons with a realized household interview are listed.

These changes imply that the geographical areas to which the county and municipality codes refer changed over time. For interpretation of macro-indicators at the county or municipal level, this hinders a continuous analysis of the characteristics of the original areas over time. A further obstacle for longitudinal analyses is that the reforms of county and municipality boundaries in the different states of East Germany began and ended at different times. It is therefore impossible to assign each SOEP household to the county or municipality to which it officially belonged at the time of the interview. SOEP households are thus assigned the county and municipality identifiers that were applicable at official reference dates.

The official reassignments of county and municipality boundaries are important for analyses of SOEP geocode data because the limitations that apply to the use of regional policy region data may be similar to those for the county and municipal data. Errors in assigning the accurate county code for the exact point in time of the interview may translate into errors in assigning the correct ROR identification number.

However, since regional policy regions refer to much larger geographical areas than counties, it is likely that this problem is negligible for analyses at regional policy region level. Researchers using these data for households living in East Germany in the 1993-1996 period should be aware of this problem and should be particularly cautious when analysing their results. None of these caveats apply to analyses of SOEP geocode data for West Germany or for analyses of East Germany after 1996. See Blach und Jonetzko (1999) for further information on the reforms and their implications for longitudinal analyses of geo-referenced data.

## **2.2 Implications of ROR readjustments in 1996**

When analysing SOEP data at the level of regional policy regions (ROR) for a number of years it also has to be considered that readjustments of these areas have taken place. ROR in 1985-1995 refer to different areas than regional policy regions from 1996 onwards. The total number of areas remained constant (97 regions). This makes it difficult to conduct longitudinal analyses covering both periods.

The readjustment of ROR was triggered by the reform of county and municipal boundaries in East Germany during the 1993-1996 period. Most of the boundaries of ROR remained as they had been. A substantial change, however, took place in the federal states of Berlin and Brandenburg. The readjustment of ROR boundaries in East Germany also was taken as a chance to

review the boundaries of ROR in West Germany. Economic interlinkages within regions may have changed, justifying the realignment of area boundaries. In West Germany, ROR changed in the federal state of Lower Saxony, in the outskirts of the federal city states of Bremen and Hamburg, in the region of Rhine-Taunus and the administrative region of Kassel (in the federal state of Hesse), and in the regions of Rhine-Ruhr and Cologne (in the federal state of North Rhine-Westphalia).

Comparing area characteristics before and after 1996 is therefore not possible for all the ROR. Analyses within these two periods separately, pre-1996 and post-1996, are not affected by the readjustment of ROR. For further information on the readjustment, see Böltken (1996).

The SOEP geocode data CD-ROM is organised according to unique household identifiers. For all years from the 1985 wave of the survey onwards, it contains the household identifiers and variables that indicate which ROR the household lives in at the time of the interview. In addition, the disk contains 79 regional indicators at the ROR level for the years 1985 to 1994 for West Germany. Regional indicators for the period after 1994 are not distributed with SOEP. They can be obtained directly from BBR. Since the mid-1990s, BBR has distributed its regional indicators in electronic form. These CD-ROMS (INKAR-CD-ROMS, “*Indikatoren und Karten zur Raumentwicklung*“) allow researchers to compile their own tables of BBR macro-indicators and to export files (all standard formats supported). For the first time, INKAR 2004 contains time series data on regional indicators.

There are a number of studies that draw on data derived from SOEP geocode. They can be found in the SOEP-lit databank at <http://panel.gsoep.de/soeplit/><sup>5</sup>.

---

<sup>5</sup> SOEPlit includes all SOEP-based publications that have been reported to the SOEP team.

### **3 SOEP data at the county level**

We refer to counties as the sum off all counties and county-free cities. Boundaries of county-free cities (kreisfreie Städte) are identical to the boundary of the municipality (Gemeinde) in which they fall. Boundaries of counties and county-free cities do not cross the boundaries of federal states. Information at county level is available for all waves of SOEP from 1985 onwards. Overall, the number of counties (Kreise and kreisfreie Städte) in Germany currently amounts to 439 (see map in Appendix 2). As is indicated in Table 2, most of the German counties are represented in SOEP. Further information on the regional clustering of SOEP respondents at the level of counties can be obtained from the maps reported in Appendices 4-6.

Users who want to analyse SOEP county-level data can draw on a dataset containing the unique household identifiers of the SOEP households, the identification number of the county in which the SOEP household lives, and selected macro- indicators at this geographical level. The county-level dataset does not only contain the original county code but also a recoded county code that facilitates longitudinal analysis with SOEP data at the county level. In addition, the dataset contains selected geographic information for the years 1985 to 2004 and selected socio-economic indicators for the years 1995 to 2002.

#### **3.1 Information on recoded county codes**

The SOEP county level dataset contains the original county codes provided to the SOEP group by the survey institute *Infratest Sozialforschung*. These county codes were cleaned, that is, all codes identified as containing errors were corrected. Along with the original (cleaned) county code, the data set contains a recoded county code.



Table 2

**Representation of counties (Kreise and kreisfreie Städte) in GSOEP. Number of observations, 1985-2005**

<b>1985-1995</b>	<b>1985</b>	<b>1986</b>	<b>1987</b>	<b>1988</b>	<b>1989</b>	<b>1990</b>	<b>1991</b>	<b>1992</b>	<b>1993</b>	<b>1994</b>	<b>1995</b>
<b>Number of counties</b>	278	287	291	291	291	456	459	465	470	476	429
<b>Number of households in county</b>											
<i>mean</i>	44.9	42.6	41.6	39.5	38.7	35.8	34.6	34.8	33.9	32.4	35.6
<i>maximum</i>	204	200	198	189	187	182	179	177	174	169	197
<i>minimum</i>	1	1	1	1	1	1	1	1	1	1	1
<b>Number of individuals in county</b>											
<i>maximum</i>	388	363	364	338	339	339	339	327	307	291	348
<i>minimum</i>	1	1	1	1	1	1	1	1	1	1	1
<b>1996-2005</b>	<b>1996</b>	<b>1997</b>	<b>1998</b>	<b>1999</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	
<b>Number of counties</b>	430	422	425	424	435	436	437	436	438	437	
<b>Number of households in county</b>											
<i>mean</i>	35.1	33.8	37.6	35.6	58.9	54.6	58.5	56.4	55	53.6	
<i>maximum</i>	191	173	199	187	298	281	293	289	280	266	
<i>minimum</i>	1	1	1	1	1	1	1	1	1	1	
<b>Number of individuals in county</b>											
<i>maximum</i>	330	297	326	308	470	454	488	468	453	426	
<i>minimum</i>	1	1	1	1	1	1	1	1	1	1	

Source: SOEP 21. Authors' calculations.

Notes: Only households/persons with a realized household interview are listed.

Recoded county codes have been assigned in order to enable the matching of time series data at county level with multiple waves of SOEP, i.e., stretching over longer periods in which county reassignments have taken place. Recoded county codes differ from original county codes assigned in SOEP when county boundaries or identifiers have changed over time. This is particularly the case for the years prior to 1997 given the extensive redistricting at the county level carried out in East Germany.<sup>6</sup> Redistricting and renaming of counties make it impossible to integrate information from external sources that relates to different county boundaries. The respective county codes will be incompatible. To avoid this problem, East German county codes (and other codes that have changed over time, if few) are recoded to correspond with current county codes.

From the year 1997 onwards, original and recoded county codes are virtually identical. As recoding renders SOEP county codes compatible with county codes from external data sources, the recoded county codes should be used to match external macro indicators with SOEP.

The data set also contains a variable with the county name assigned on the basis of information derived from *Statistik regional*.<sup>7</sup>

### **3.2 Additional information at the county level**

SOEP has been matched on the basis of the recoded county codes with selected geographically referenced information relating to the counties in which SOEP respondents live. This allows researchers conducting SOEP analyses at the county level to familiarise themselves with the linked macro-micro data and also some basic information on the counties themselves.

The indicators supplied by SOEP include data on the total surface area, population density, income, number of employed persons making compulsory social insurance contributions, and unemployment. Researchers may match further information at the county level.

While most of the indicators supplied with SOEP are official data, some information at this level has been derived solely for SOEP. This is true for information on the longitude and latitude of the county in which the SOEP household lives. Longitude and latitude have been

---

<sup>6</sup> See also, e.g., Blach und Jonetzko (1999).

assigned (with only a few exceptions) based on the centroid of the respective county. This was established using the Microsoft program package MS Autoroute.<sup>8</sup> Information on longitude and latitude may be used to calculate the distances between two counties.

Table 3 reports summary statistics of select indicators at the county level matched with SOEP.

**Table 3**  
**Selected indicators at county level in SOEP, 2002**

<i>Indicator</i>	<i>Unit</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>Minimum</i>	<i>Maximum</i>
Area Size	<i>km<sup>2</sup></i>	808.76	590.13	35.63	3,058.23
Population	<i>individuals</i>	435,306.5	678,949.8	35,846	3,392,425
German population	<i>individuals</i>	382,168	584,014.4	34,151	2,947,651
Non-German population	<i>individuals</i>	53,138.47	98,141.85	665	444,774
Longitude	<i>degree</i>	9.838	2.216	6.090	14.990
Latitude	<i>degree</i>	50.867	1.659	47.500	54.790

Source: SOEP 21. Authors' calculations.

To date, several studies have been conducted using the county code information. The following list gives a brief overview on selected studies carried out using SOEP county codes:

Büchel, Felix und C. Katharina Spieß (2002): Müttererwerbstätigkeit und Kindertageseinrichtungen - neue Ergebnisse zu einem bekannten Zusammenhang, in: *Vierteljahrshefte zur Wirtschaftsfor-*  
*schung*, 2002 (71), 96-114).

Hank, Karsten (2003): 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 (1), 3-21.

Hank, Karsten, Michaela Kreyenfeld und C. Katharina Spieß (2004): Kinderbetreuung und Fertilität in Deutschland, in: *Zeitschrift für Soziologie*, 33 (3), 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).

Jürges, Hendrik (2005): The Geographic Mobility of Dual-Earner Couples: Does Gender Ideology Matter?, *DIW-Discussion Paper* 474.

---

<sup>7</sup> See also *Statistik regional, Daten und Informationen der Statistischen Ämter des Bundes und der Länder*, 1998 to 2004 edition (on CD-ROM).

<sup>8</sup> Compare Microsoft AutoRoute 2005 (on CD-ROM).

Rehdanz, K. and Maddison, D. (2005): Der Wert des Klimas für Haushalte in Deutschland, erscheint in: G. Grözingen/W. Matiaske (Hrsg.): *Deutschland regional: Sozialwissenschaftliche Daten im Forschungsverbund*.

Wrohlich, Katharina (2006): Child Care Costs and Mothers' Labor Supply: An Empirical Analysis for Germany, in: Applied Economics (forthcoming).

## **4 SOEP data at the zip-code level**

For all waves including and following the 1993 wave of the study, SOEP records the five-digit zip-code area in which respondents live at the time of the (household) interview. Before we present raw frequencies of zip-code areas represented in SOEP (in both the cross-sectional and longitudinal perspective) and point out potentials and limitations for socio-scientific analyses, we provide researchers interested in working with this feature of the panel study with some elementary information on the geography of zip-code areas.

### **4.1 Geographical reference of zip codes**

In Germany zip-code areas are defined by the postal service Deutsche Post AG to optimise mail delivery. In 1961, both West and East Germany introduced a four-digit system. Both of these systems were merged into one system after German reunification because 802 of the former zip codes were identical. The current five-digit system for the whole of Germany replaced the two previous systems on the 1 July 1993.

Deutsche Post AG assigned zip codes without consideration of political or administrative boundaries but exclusively with respect to topography and operating processes within their company (at the local level). It follows that the zip-code system is a dynamic one, i.e., Deutsche Post AG can add new zip codes or drop or reassign existing ones when this promises to expedite mail delivery. Zip-code reassignments can be triggered, e.g., by migration of households into or out of existing zip-code areas or by Deutsche Post AG deciding to reduce the number of delivery people working in a particular area. In both cases the workload for delivery people will change, and if this makes the existing zip-code structure inefficient at the local level, a re-organisation will take place. Each delivery person carries mail to approximately 1,300 individuals or 660 households, i.e., when an analysis of SOEP at the zip-code level is undertaken, the matched neighbourhood indicators refer to (estimated) average characteristics of approximately 26,000 people, i.e., 13,200 households (figures as of 2001, Source: GfK).

Box 1

**Zip codes – the meaning of the digits**

---

The German zip-code system is comprised of ten different master regions (*Leitzonen*) that are identifiable by the first digit of the code. Each of these master regions is further divided into up to ten regions (*Briefregionen*). **The first two digits** of the zip-code level are thus very unlikely to change, so knowing the first two digits can enable one to roughly identify where the location. For instance, the zip-code 01xxx relates to an area in Dresden, 10xxx to one in central Berlin, and 99xxx to an area in the city of Erfurt. There are a total of 83 *Briefregionen* given that some of the larger master regions such as Berlin or Munich could not be covered by just a second digit.

**The third to fifth digits** of the zip-code are assigned by the local post office, and they relate to entry types (*Eingangseinheiten*) of the zip-code. There are three different entry types. The first type is a mailbox. The second type is a company or other organisation that receives a great deal of mail. The third entry type relates to real areas and identifies a mail delivery region (*Zustellbezirksgruppe*) covered by an average of 20 postmen. The third type of zip code is the geographical unit relevant for socio-scientific analyses of area effects. These real locations covered by a single zip code usually are confined areas in a large city, or a cluster of small villages. In the latter case, zip-code areas can consist of quite a large number of villages: using the online zip-code search engine provided by Deutsche Post (2004), we identified zip-code areas containing as many as 41 small villages.

Within a local delivery system, entry types systematically order zip codes. However, local delivery systems do not apply the same assignment system. In other words, from the macro perspective, zip codes appear to be rather randomly assigned, that is, beyond the first two digits.

---

## 4.2 SOEP sample distributions by zip-code areas

To give the reader some idea about how zip-code areas are represented in SOEP, Table 4 provides descriptive information on the total number of zip-code areas in Germany at three points in time and also indicates how many of these areas contain SOEP households (expressed as percent of all zip-code areas). We distinguish between the ten zip-code master regions that exist in Germany (see map in Appendix 3). Figures are provided for three points in time (1993, 1998 and 2004).

The results show that about 30 percent of the zip-code areas are represented in the SOEP. Over time, this proportion has increased. In 1993, SOEP households were found in 26 percent of the zip-code areas. The numbers for 1998 and 2004 are 28 and 32 percent, respectively.

These increases are due to migration of households within Germany. The large increase between 1998 and 2004 is also due to the expansion of the SOEP sample in 2000.

Table 4

**Representation of unique zip-code areas in SOEP, 1993, 1998 and 2004**

zip-code region	Total number of zip-code areas in Germany			Zip-code areas represented in SOEP			
				% of all zip-codes			
	1993 <sup>1</sup>	1998 <sup>1</sup>	2004 <sup>2</sup>	1993	1998	2004	$\bar{\theta}$ 1993/1998/2004
0	675	674	658	27.0	37.1	50.6	38.2
1	541	540	538	36.4	46.7	64.7	49.3
2	987	984	984	20.9	27.6	38.8	29.1
3	813	814	814	29.5	36.0	47.8	37.8
4	550	550	550	49.8	56.2	72.2	59.4
5	712	712	712	32.2	40.2	51.0	41.1
6	687	690	691	30.6	36.2	48.9	38.6
7	981	982	981	23.2	27.3	39.3	29.9
8	1140	1137	1134	19.2	22.6	32.9	24.9
9	1186	1187	1187	14.3	18.2	27.6	20.0
Total	8274	8270	8249	26.0	27.9	32.5	28.8

Source: SOEP 21. Authors' calculations.

Notes: 1) As of 31 December of the respective year. Source: Infas Geodata.

2) As of 31 December of the respective year. Source: Microm

There is substantial variation in the extent to which the zip-code regions are represented, both within regions over time and across regions. Zip-code region 6 has the highest percentage of zip-code areas (60 percent on average) and region 9 the lowest (20 percent). Over time, we see increases in the percentages of zip-code areas in all zip-code regions, owing to the addition of two new SOEP samples (Sample F in 2000 and Sample G in 2002), and also to relocations of individuals and households. The biggest increase took place in zip-code region 1 (from 36 percent in 1993 to 64 percent in 2004).

Table 5 shows the number of zip-code areas represented in SOEP, differentiated by zip-code master region. Figures are presented for all years from 1993 to 2004.

**Table 5**  
**Number of zip-codes represented in SOEP differentiated by zip-code region, 1993-2004**

zip-code region	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
0	181	190	207	230	235	250	251	294	302	311	311	333
1	197	203	222	232	247	252	253	286	306	318	329	348
2	206	223	248	241	248	272	279	329	338	348	363	382
3	240	264	283	289	294	293	295	347	365	365	370	389
4	274	285	296	290	288	309	306	350	360	377	374	397
5	229	243	261	266	270	286	293	326	332	332	336	363
6	210	223	237	241	239	250	258	295	306	306	313	338
7	228	253	263	262	268	268	263	316	342	343	343	386
8	219	242	253	244	247	257	264	312	331	339	338	373
9	170	182	195	203	212	216	220	267	275	296	300	328
Total	2155	2308	2465	2498	2548	2653	2682	3122	3257	3335	3378	3637

Source: SOEP 21. Authors' calculations.

For analyses of SOEP data on the zip-code level, researchers may be interested in the number of households found in each zip-code area. It is not possible to provide this information for individual zip-code areas for data protection reasons: such information could theoretically enable people to identify SOEP respondents (for instance, if many households from a very tiny zip-code area are represented). We can, however, disclose this information on the level of zip-code master regions. Table 6 presents the results.

It can be seen that the SOEP has a sizeable number of households in each zip-code region in each year of the survey. In no wave of the survey do we observe less than 500 households per zip-code master region. The smallest number of households is in zip-code region 9 in 1993 (n=532), and most households are found in zip-code region 3 in 2001 (n=1,680). In about 50 percent of the years and regions, we observe more than 1,000 households.



**Table 6**  
**Number of households in SOEP differentiated by zip-code regions, 1993-2004**

zip-code region	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
0	1058	1019	1016	1002	988	1083	1057	1627	1620	1508	1474	1491
1	843	810	824	812	796	859	833	1307	1303	1259	1211	1251
2	544	566	584	558	554	692	687	1282	1275	1194	1136	1208
3	881	923	971	950	938	1014	995	1672	1680	1581	1490	1553
4	746	765	788	772	761	929	915	1678	1664	1573	1462	1546
5	687	736	777	751	748	856	850	1471	1472	1360	1294	1371
6	633	657	724	706	693	768	764	1304	1298	1224	1162	1220
7	769	788	811	792	775	872	855	1372	1385	1253	1182	1270
8	612	619	644	617	601	689	672	1141	1149	1090	1026	1110
9	532	547	578	568	566	661	670	1205	1195	1105	1039	1070
Total	7306	7430	7717	7528	7420	8423	8298	14059	14041	13147	12477	13090

Source: SOEP 21. Authors' calculations.

Table 7 presents the respective figures on the level of individuals interviewed in the years 1993 to 2004.

**Table 7**  
**Number of individuals in SOEP differentiated by zip-code regions, 1993-2004**

zip-code region	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
0	1920	1904	1872	1860	1839	1953	1917	2985	2755	2625	2565	2598
1	1408	1396	1413	1374	1327	1384	1342	2120	2004	1918	1913	2003
2	926	950	958	973	955	1152	1099	2123	1910	1879	1804	1988
3	1611	1663	1736	1707	1698	1809	1739	3016	2777	2629	2513	2693
4	1351	1375	1399	1387	1345	1594	1516	2911	2645	2480	2420	2621
5	1258	1311	1407	1390	1369	1532	1443	2601	2337	2222	2137	2261
6	1196	1247	1321	1259	1231	1351	1303	2281	2082	1986	1924	2010
7	1449	1485	1491	1432	1444	1529	1432	2446	2153	2056	1973	2183
8	1057	1048	1080	1074	1045	1136	1105	1926	1763	1675	1601	1812
9	1003	1038	1091	1054	1029	1230	1189	2177	1925	1751	1729	1843
Total	13179	13417	13768	13510	13282	14670	14085	24586	22351	21221	20579	22012

Source: SOEP 21. Authors' calculations.

As a longitudinal study, SOEP also allows the analysis of changes over time. In terms of neighbourhood characteristics, these changes may be in the characteristics of the neighbourhood. But they may also be demographic changes in neighbourhoods over time, for example, when individuals relocate.

Table 8 indicates how many of the SOEP households live in different zip-code areas in year  $t$  and  $t+1$  and, secondly, in year  $t$  and  $t+5$ . Figures are presented for all years from 1993 to 2003 (changes from  $t$  to  $t+5$  for 1993 to 1999 only).

### **4.3 Potential for Analyses of SOEP data at the zip-code level**

#### ***Neighbourhood indicators***

Matching SOEP data with micro-geographical information at the zip-code level would allow researchers to investigate whether neighbourhood context effects exist on a sample representative of the entire German population. Neighbourhood indicators may be gathered from a number of different sources. On the one hand, the Statistical Offices of German cities and towns may be able to provide data on key socio-demographic characteristics of the population in their area at the zip-code level. Indicators include the number of residents in the area, the number of foreigners, and the age and family structures in the area. Not all indicators will be available in all cities, and data are not usually provided free of charge (concessions may be made for researchers). It is also possible to use official data at other geographical scales and link them with SOEP data at the zip-code level. To do so, researchers need to know which zip-code area corresponds to the other geographical area in question.

On the other hand, suppliers of micro-geographical data (such as Postdirekt, Microm or GfK to name but a few) offer a wide range of indicators at a spatial scale corresponding to the zip-code level. Most of these purchasable indicators are estimates of neighbourhood characteristics that might be affected by measurement error, in some cases significantly.

Table 8

**Number of households in SOEP who still lived in the same zip-code area in the subsequent wave and five years later**

(unweighted frequencies)

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Number of households observed in t and t+1	6220	6390	6566	6496	6301	7052	6883	11530	10936	11545	10431
Number of households in different zip-code area in t and t+1	245	276	432	308	327	334	320	491	444	457	417
<i>% of all households</i>	3.9	4.3	6.6	4.7	5.2	4.7	4.6	4.3	4.1	4.0	4.0
Number of households observed in t and t+5	5155	5285	5327	5254	5178	5780	5729		n.a.		
Number of households in different zip-code area in t and t+5 <sup>1</sup>	802	891	951	924	894	933	910		n.a.		
<i>% of all households</i>	15.6	16.9	17.9	17.6	17.3	16.1	15.9		n.a.		

Source: SOEP 21. Authors' calculations.

Notes: 1) Intermediate changes not considered.

Suppliers of geodata build their estimation models on data gathered through telephone interviews, local statistics, mail-order data and the like, and are reluctant to provide more detailed information on their estimation procedures. Their data sources may not fulfil academic standards (in terms of response rates, sampling issues, response biases). Indicators include, for instance, the average disposable income available in the neighbourhood, the percentage of foreigners in the zip-code area, (3) the unemployment rate, (4) the social composition of the neighbourhood in form of a milieu typology, and also (5) the balance of inward and outward migration.

### ***Research questions***

A number of studies in Germany have focussed on the problem of increasing local segregation and its consequences (e.g. Häußermann & Kapphann 1999; Alisch & Dangschat 1993; Farwick 2001). All of these are case studies focussing on one or two German cities. Once SOEP data have been matched with neighbourhood indicators, however, analyses of neighbourhood effects on individual-level outcomes are possible for Germany as a whole as well.

An example of the kind of research that could be undertaken on the basis of the SOEP is Buck's study "Identifying Neighbourhood Effects on Social Exclusion" (Buck 2001) on the basis of the British Household Panel Study (BHPS). Further examples of studies using the BHPS include McCulloch's study of ward-level deprivation and a number of different socio-economic outcomes (McCulloch 2001) and the work of Propper et al., who analyse the impact of neighbourhood socio-economic characteristics on mental health outcomes (Propper, Jones et al. 2004). Drever uses SOEP data matched with an indicator of the percentage of foreigners at the zip-code level for a number of German cities to investigate whether there are negative effects of living in ethnic neighbourhoods for immigrants in Germany (Drever 2004). Knies matches SOEP data at zip-code level with an indicator of the average neighbour's income to establish whether neighbourhood incomes represent a negative externality for people's life satisfaction in Germany (Knies 2005; Knies 2006).

### ***A cautionary note***

Analysis of data at the zip-code level as a proxy for spatially defined neighbourhoods has downsides, mainly owing to the fact that the system is dynamic and that changes in zip codes

are (almost) intractable. If changes occurred over time in the characteristics of a zip-code area's population, these changes may have been caused by inward and outward migration of people to and from zip-code areas with fixed boundaries. However, they may also have changed due to a change in the boundaries of the zip-code area. In this case the characteristics of the original neighbourhood population may not have changed, but since more people are considered neighbours, the character of the neighbourhood appears to have changed. Finally, both the boundaries and the population of a zip-code area may be constant over time but the socio-economic characteristics of the population may have changed. Analytically, these changes cannot be disentangled because it is not clear which reassignments were undertaken by Deutsche Post and at what time.

It is difficult to think of neighbourhoods as places where social interactions take place in the form of reciprocal behaviour and where people share common points of reference given the large number of people considered neighbours when 'neighbourhood' is operationalised at the zip-code level. However, according to figures provided by market researchers in their micro-geographical data, the actual number of people in zip-code areas varies widely and is generally higher in rural areas than in metropolitan areas where households live in higher density. In terms of space, urban postal areas may be assumed to approximate people's idea of 'neighbourhood' better than rural postal areas.

Selected papers using SOEP zip-code data are listed below:

- Anger, Silke (2005): Unpaid Overtime in Germany: Differences Between East and West, *Journal of Applied Social Science Studies* (Schmollers Jahrbuch), 125 (1), 17-21.
- Anger, Silke (2005): Unbezahlte Überstunden und regionale Arbeitslosigkeit. In: Grözinger, Gerd, and Wenzel Matiaske (Hrsg.): *Deutschland regional – Sozialwissenschaftliche Daten im Forschungsverbund*, München/Mering: Rainer Hampp, 227-245.
- Drever, Anita I. (2004): Separate Spaces, Separate Outcomes? Neighbourhood Impacts on Minorities in Germany, in: *Journal of Urban Studies*, (41: 1423-1439).
- Knies, Gundi (2005): "Keeping up with the Schmidts: Do richer neighbours make people unhappier?", mimeo.

## 5 Further information at the neighbourhood level

The SOEP also collects a considerable amount of data on the residential area where individuals live. Some of these data are available for all years of the survey, and some are collected in those waves of the panel focusing on neighbourhood infrastructure and social networks. The SOEP waves with this focus are 1986, 1994, 1999, and 2004. SOEP provides researchers with indicators on the built environment, local infrastructure, the natural environment, and also the social environment.

### 5.1 Indicators of the built environment

Basic characteristics of respondents' residential living areas are provided in the \$hbrutto data file. The relevant indicators are "\$wum1" and "\$wum3", which are generated by the SOEP group on the basis of information households provide in the wave-specific household questionnaire if they relocated since the previous wave (and when a new household enters the survey). If no change of residence took place, the variables "\$wum1" and "\$wum3" contain the information provided in the survey subsequent to the last move.<sup>9</sup>

The indicator "\$wum1" classifies the respondent's home as [1] farm house, [2] single occupancy (1-2 family house), [3] single occupancy (1-2 family house) in a row house, [4] apartment flat in a 3-4 unit apartment building, [5] apartment flat in a 5-8 unit apartment building, [6] apartment flat in a +9 unit apartment building, [7] high-rise, or [8] other building.

The variable "\$wum3" tells whether the area they now live in is [1] a residential area with predominantly old houses, [2] a residential area with predominantly new houses, [3] a mixed area with shops and houses, [4] a predominantly industrial area, or [5] another kind of neighbourhood.

The information contained in the variables "wum1" and "wum3" may be used independently or combined, for example, with area descriptions such as city size (provided in the BOUST-EDT instrument of SOEP, see Section 1 above), or compressed into a community typology.<sup>10</sup>

---

<sup>9</sup> In borderline cases, this may mean that a wave-specific area description is inadequate—for example when large-scale redevelopment has taken place since the move, placing all residential areas in a mixed housing and shopping area. We cannot avoid this problem due to lack of further information.

<sup>10</sup> This has been done by Peter Bartelheimer of the University of Göttingen (Soziologisches Forschungsinstitut Göttingen, SOFI, Georg-August-Universität).

## 5.2 Indicators of the local infrastructure and the natural environment

To allow researchers to get a clearer picture of the particular environments individuals are confronted with, SOEP provides indicators on the availability of local public amenities and also on the quality of the natural environment. These indicators can be derived from the SOEP household questionnaires that have a special focus on neighbourhood infrastructure and social networks. This is the case for the 1986, 1994, 1999 and 2004 waves of the survey.

The head of household provides information on how long it takes to reach a number of basic amenities on foot. The list includes (1) basic infrastructure such as doctor, shops, public transport and banks or automatic teller machines, (2) institutions catering to particular age groups such as kindergartens, primary schools, youth clubs and senior centres, (3) facilities for recreational and leisure activities such as sports clubs, pubs, and parks. Furthermore, the head of household is asked whether (4) the levels of environmental strain (air or noise pollution) or the lack of parks or green space in the neighbourhood are disconcerting.

Tables 9.1 to 9.4 present the raw frequencies of all the characteristics for all four years. The figures document that there is considerable variation in the availability of these amenities and/or levels of concern about the environmental strain, both per indicator and over time.

Table 9.1

**Number of SOEP households differentiated by how long it takes to walk to local basic amenities. 1986, 1994, 1999 and 2004**

	Shops				Bank			
	1986	1994	1999	2004	1986	1994	1999	2004
less than 10 min.	3579	4333	4477	6582	3260	3742	3899	5722
10 - 20 min.	1148	1771	2024	3291	1349	2008	2284	3724
more than 20 min.	222	352	398	763	289	531	601	1038
Not available, not accessible on foot	116	290	435	1096	164	460	535	1234
	Doctor				Public Transport			
	1986	1994	1999	2004	1986	1994	1999	2004
less than 10 min.	2434	2832	2872	4297	4318	5592	6055	9925
10 - 20 min.	1450	2139	2423	3627	596	992	1097	1532
more than 20 min.	636	880	989	1761	79	87	114	156
Not available, not accessible on foot	529	878	1019	2007	56	60	47	104

Source: SOEP 21. Authors' calculations.

Table 9.2

**Number of SOEP households differentiated by how long it takes to walk to local facilities for particular age groups, 1986, 1994, 1999 and 2004**

	Kindergarten				Primary School			
	1986	1994	1999	2004	1986	1994	1999	2004
less than 10 min.	2751	3173	3388	5379	2566	2883	3065	4635
10 - 20 min.	1275	1917	2095	3256	1463	2104	2243	3554
more than 20 min.	335	451	525	985	418	564	632	1193
Not available, not accessible on foot	250	403	476	955	248	539	677	1355
	Youth Club				Old-age Facility			
	1986	1994	1999	2004	1986	1994	1999	2004
less than 10 min.	-	1814	2040	2922	-	1668	1736	2707
10 - 20 min.	-	1901	2134	3274	-	1693	1991	3002
more than 20 min.	-	776	909	1684	-	898	1007	1811
Not available, not accessible on foot	-	1377	1292	2339	-	1684	1791	3009

Source: SOEP 21. Authors' calculations.

Table 9.3

**Number of SOEP households differentiated by how long it takes to walk to local facilities for leisure activities, 1986, 1994, 1999 and 2004**

	Pub / Restaurants				Parks/ Green Space			
	1986	1994	1999	2004	1986	1994	1999	2004
less than 10 min.	-	4513	4758	7224	3029	3530	4142	6717
10 - 20 min.	-	1696	1956	3293	1166	1666	1694	2713
more than 20 min.	-	323	344	631	350	559	598	820
Not available, not accessible on foot	-	162	219	521	427	823	734	1233
	Sports Centre / Playing Yard							
	1986	1994	1999	2004				
less than 10 min.	2223	2393	2864	4604				
10 - 20 min.	1824	2436	2545	4171				
more than 20 min.	662	1040	1086	1646				
Not available, not accessible on foot	252	630	587	970				

Source: SOEP 21. Authors' calculations.



Table 9.4

**Number of SOEP households differentiated by how much they are bothered by noise pollution, pollution and lack of access to green space, 1986, 1994, 1999 and 2004**

	Noise pollution				Pollution			
	<i>1986</i>	<i>1994</i>	<i>1999</i>	<i>2004</i>	<i>1986</i>	<i>1994</i>	<i>1999</i>	<i>2004</i>
Not At All	1894	2025	2464	4965	1893	1799	2588	5525
Slightly	1696	2579	3097	4467	1697	2744	3152	4582
Bearable	801	1357	1215	1537	857	1429	1119	1166
Strongly	487	578	419	590	457	590	377	362
Very strongly	205	230	151	187	177	210	116	86
	Lack of green space							
	<i>1986</i>	<i>1994</i>	<i>1999</i>	<i>2004</i>				
Not At All	-	4377	5081	8281				
Slightly	-	1371	1496	2486				
Bearable	-	633	553	636				
Strongly	-	278	168	228				
Very strongly	-	74	24	50				

Source: SOEP 21. Authors' calculations.

### 5.3 Indicators of the social environment

In addition to these indicators that capture the respondent's physical environment, the head of household is asked a number of questions regarding the household's social environment.

In all four waves of the survey that focus on social networks and community, the head of household was asked to indicate whether there are foreigners living in the neighbourhood. The answer categories are [1] yes, many, [2] yes, a few, [3] no, and [4] do not know. Table 10 shows the number of households that live in neighbourhoods with these characteristics.

Table 10

**Presence of foreigners in residential areas 1986, 1994, 1999 and 2004**

	Number of households (unweighted)			
	<i>1986</i>	<i>1994</i>	<i>1999</i>	<i>2004</i>
Many foreigners	1056	1283	1561	2332
Few foreigners	2053	2592	3040	5395
No foreigners	1670	2254	2164	3263
Unknown	301	633	596	762

Source: SOEP 21. Authors' calculations.

The head of household is also asked to indicate which of the following descriptions characterises the social relations between the people in their neighbourhood best: [1] people hardly know each other, [2] people sometimes talk to each other, [3] people have a relatively close relationship, or [4] peoples' relationships vary. In addition, the head of household describes his or her household's relationship to their neighbours. Answer categories range from [1] very close to [5] no contact. Finally, information is collected on whether people have close enough contacts to their neighbours to visit each other at home, and if so, with what frequency. Table 11 presents the raw frequencies for 1986, 1994, 1999 and 2004.

Table 11  
**SOEP households and social relations in the neighbourhood. 1986, 1994, 1999 and 2004**

Relationships of neighbours to each other	Number of households (unweighted)			
	1986	1994	1999	2004
Hardly know each other	-	636	670	948
Occasionally speak to each other	-	3916	4178	6852
Close bond	-	1519	1710	2750
Varied	-	688	752	1133
Own relationship to neighbours	Number of households (unweighted)			
	1986	1994	1999	2004
Very close	350	334	331	593
Close	1059	1418	1505	2595
Average	2279	3300	3620	5840
Not so close	1157	1480	1648	2409
Almost no contact	237	223	226	293
Visits with neighbours	Number of households (unweighted)			
	1986	1994	1999	2004
Yes	2954	3775	4009	6534
No	2128	2992	3339	5201
Frequency of visits with neighbours	Number of households (unweighted)			
	1986	1994	1999	2004
Almost every day	452	568	482	640
At least once per week	1190	1404	1499	2158
At least once per month	776	1012	1105	1980
Less often	529	774	896	1719

Source: SOEP 21. Authors' calculations.

## References

- Blach, Antonia und Jacek Jonetzko (1999): „Die Gebietsreform der neuen Länder: Folgen für die Laufende Raumbeobachtung 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.
- Buck, Nick (2001): "Identifying neighbourhood effects on social exclusion." *Urban Studies* 38(12): 2251-2275.
- Deutsche Post (2004): zip code search engine, [http://www.postdirekt.de/plz\\_suche/](http://www.postdirekt.de/plz_suche/).
- Drever, Anita I. (2004): "Separate spaces, separate outcomes? Neighbourhood impacts on minorities in Germany." *Urban Studies* 41(8): 1423-1439.
- Frick, Joachim R. and Jan Goebel (2005): "Regional Income Stratification in Unified Germany Using a Gini Decomposition Approach," DIW Berlin Discussion Paper 540
- Goebel, Jan (2005): "Job submission instructions for the SOEP-remote System at DIW Berlin", mimeo.
- Knies, Gundi (2005): "Keeping up with the Schmidts: Do richer neighbours make people unhappier?", mimeo.
- Knies, Gundi (2006): Trying the Neighbourhood for Size, mimeo.
- McCulloch, Andrew (2001): "Ward-level deprivation and individual social and economic outcomes in the British Household Panel Study." *Environment and Planning A* 33(4): 667-684.
- Propper, Carol, Kelvin Jones et al. (2004): "Local neighbourhood and mental health: Evidence from the UK." CMPO Working Paper Series 04(099).

## Appendix

### Appendix 1

#### Map of Federal States of Germany



#### Legend:

##### West Germany

- 0 Berlin
- 1 Schleswig-Holstein
- 2 Hamburg
- 3 Lower Saxony
- 4 Bremen
- 5 North Rhine-Westphalia
- 6 Hesse
- 7 Rhine.-Palatinate, Saarl.
- 8 Baden-Wuerttemberg
- 9 Bavaria

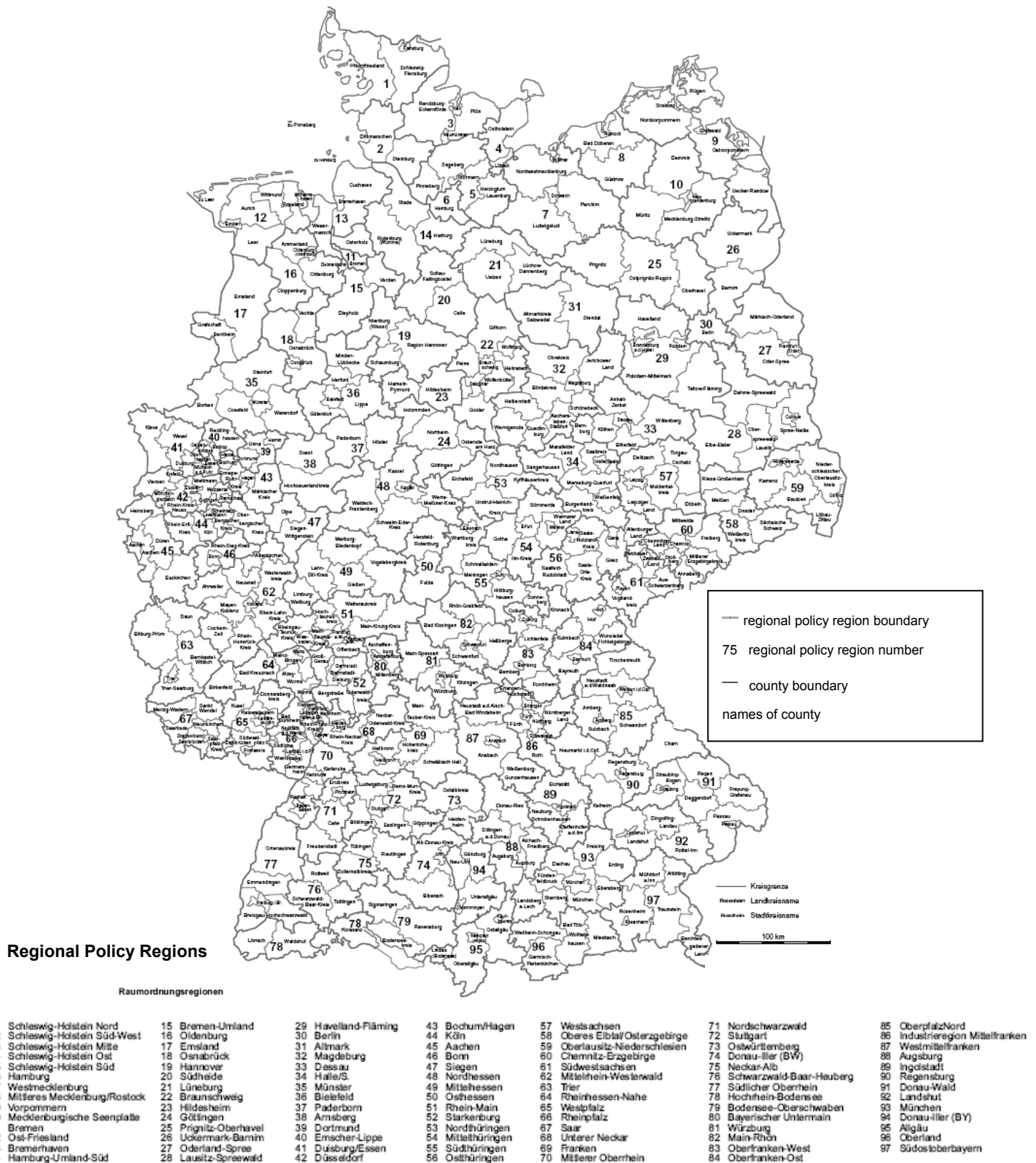
##### East Germany

- 11 East Berlin
- 12 Mecklenburg-West Pomerania
- 13 Brandenburg
- 14 Saxony-Anhalt
- 15 Thuringia
- 16 Saxony

Source: [http://europa.eu/abc/maps/members/germany\\_en.htm](http://europa.eu/abc/maps/members/germany_en.htm)

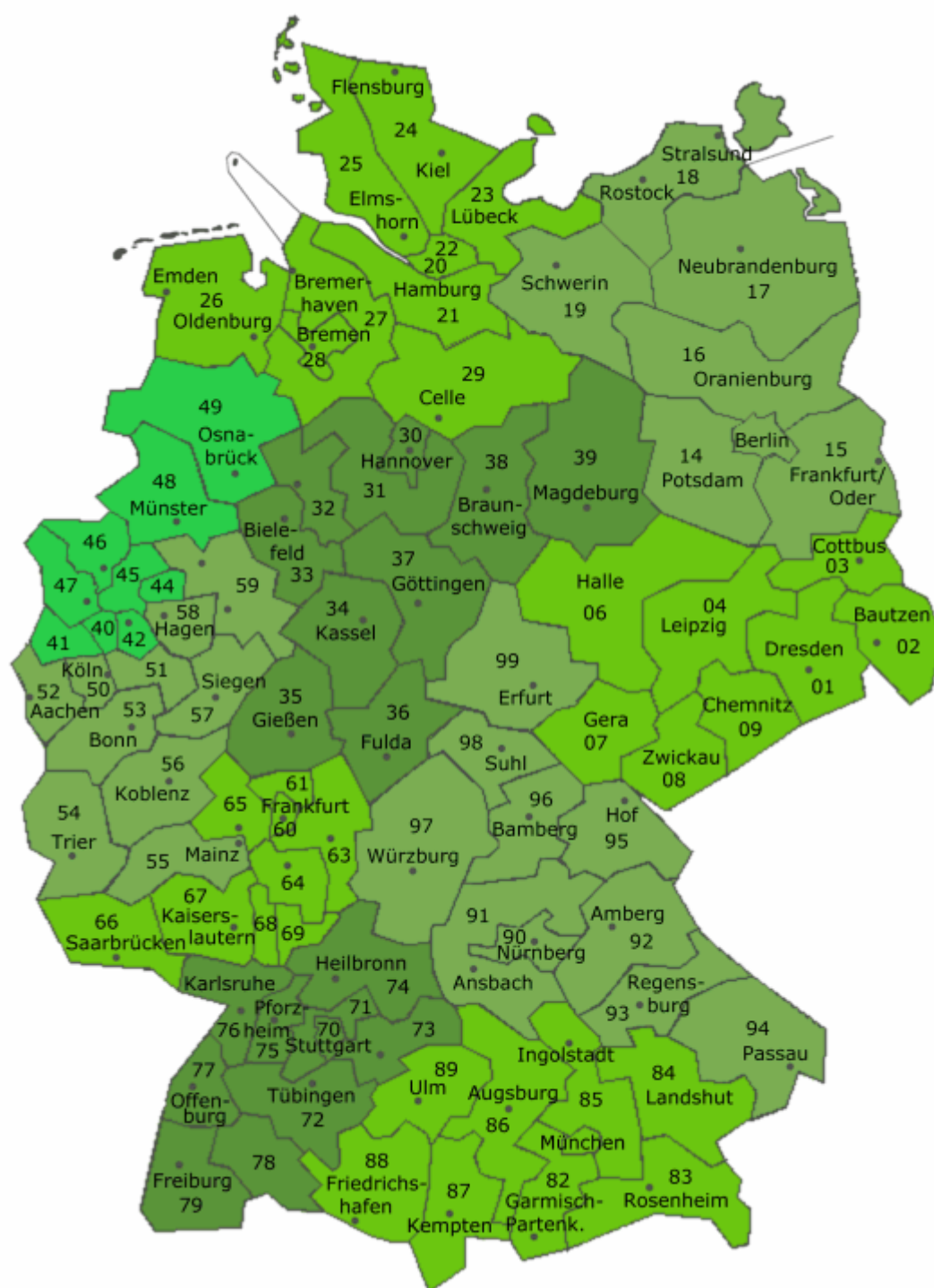
## Appendix 2

## Map of Regional Policy Regions (ROR) and Counties in Germany, 2004



## Appendix 3

## Map of zip-code regions in Germany (first digit identifies zip-code master region)

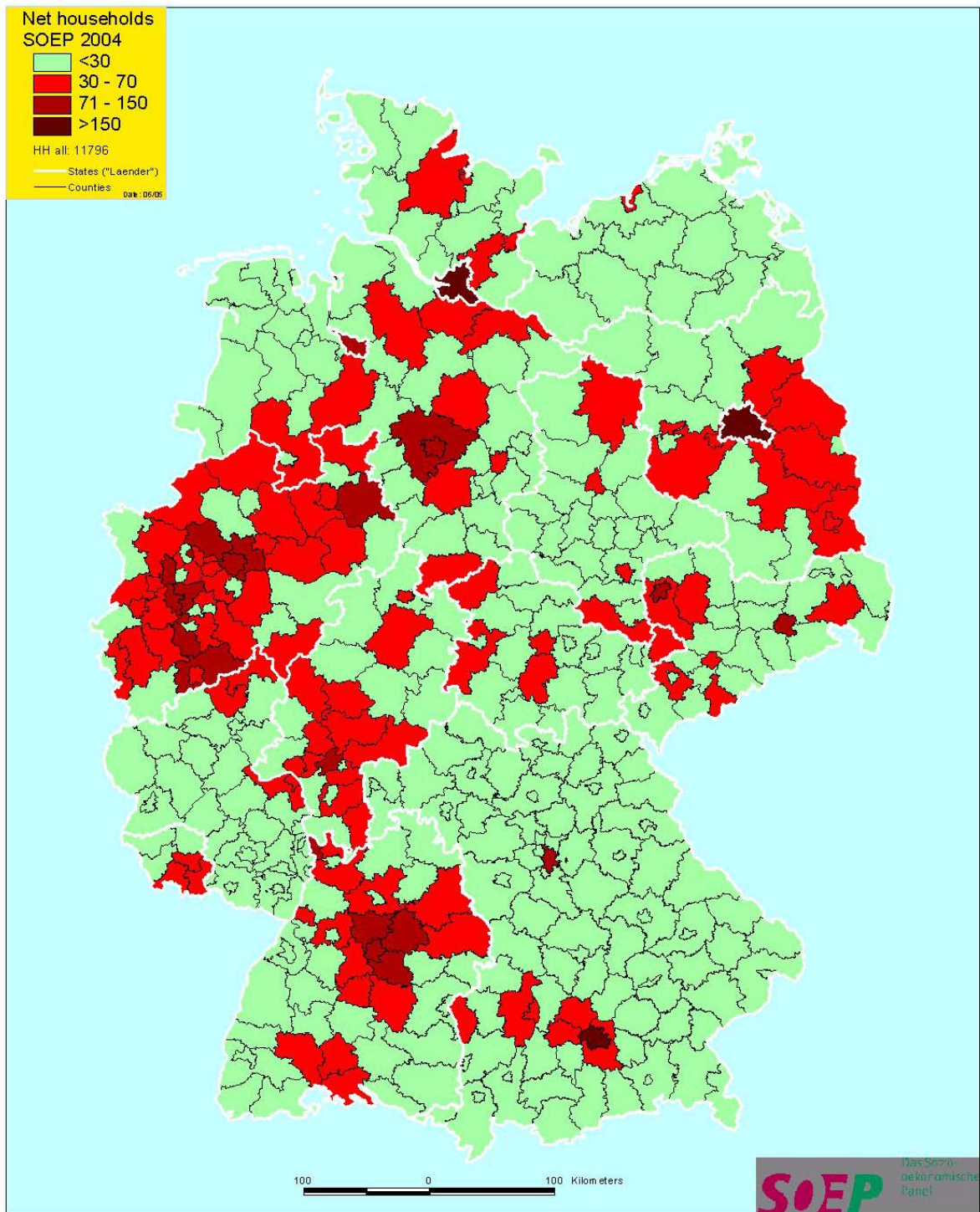


Source: Postleitzahlenkarte Deutschland. <http://www.archive.nrw.de/index.asp>



Appendix 4

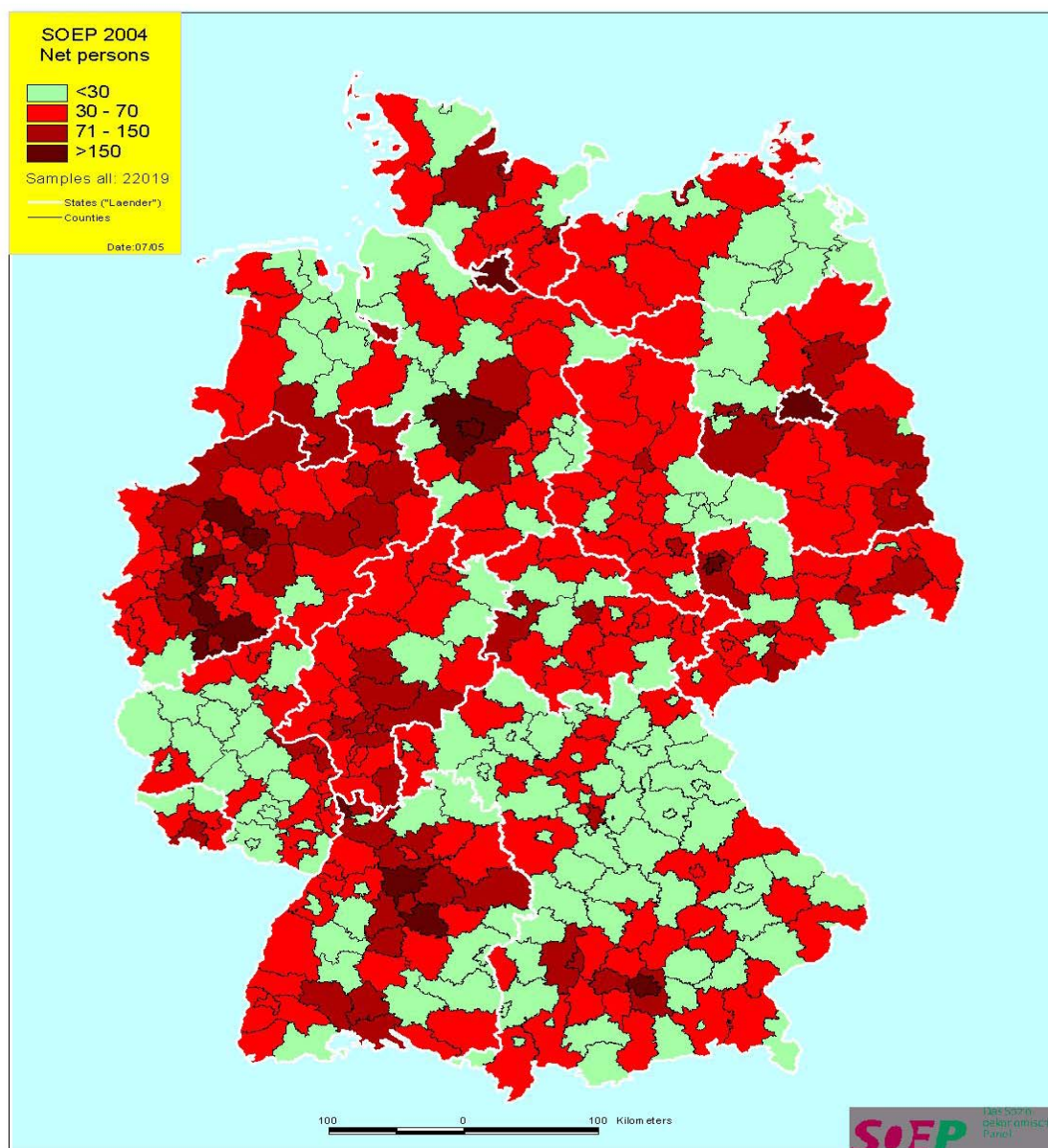
Map of counties in Germany, differentiated by number of SOEP households observed in 2004



Source: DIW Berlin, Marco Mundelius.

## Appendix 5

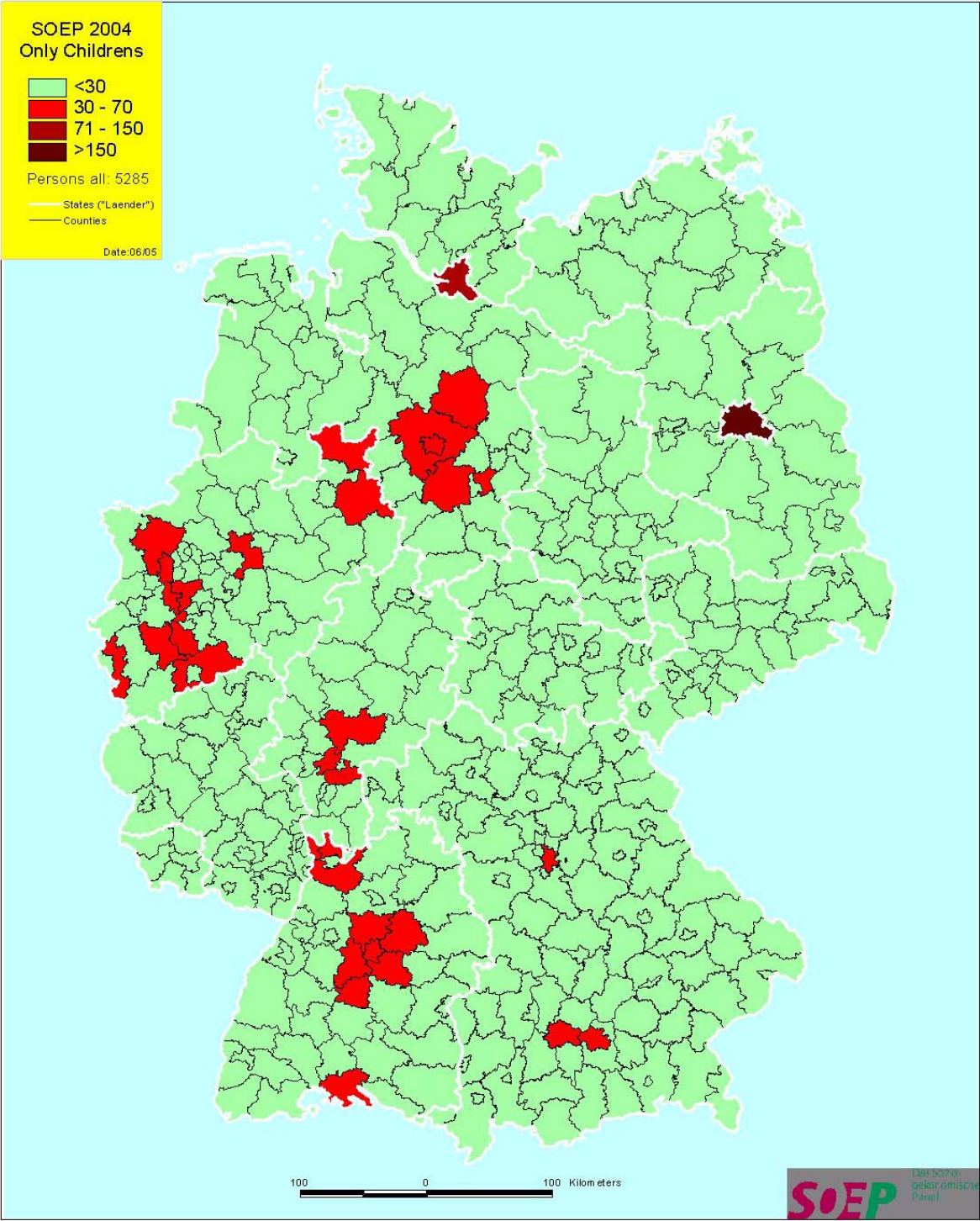
### Map of counties in Germany, differentiated by number of SOEP individuals observed in 2004



Source: DIW Berlin, Marco Mundelius.



Appendix 6  
Map of counties in Germany, differentiated by number of SOEP children observed in 2004



Source: DIW Berlin, Marco Mundelius.