

Burkhauser, Richard V.; Kreyenfeld, Michaela; Wagner, Gert G.

Article — Digitized Version

The German Socio-Economic Panel: A Representative Sample of Reunited Germany and its Parts

Vierteljahrshefte zur Wirtschaftsforschung

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Burkhauser, Richard V.; Kreyenfeld, Michaela; Wagner, Gert G. (1997) : The German Socio-Economic Panel: A Representative Sample of Reunited Germany and its Parts, Vierteljahrshefte zur Wirtschaftsforschung, ISSN 0340-1707, Duncker & Humblot, Berlin, Vol. 66, Iss. 1, pp. 7-16

This Version is available at:

<https://hdl.handle.net/10419/141155>

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.

The German Socio-Economic Panel: A Representative Sample of Reunited Germany and its Parts

By Richard V. Burkhauser, Michaela Kreyenfeld, and Gert G. Wagner

Summary

The 1995 version of the German Socio-Economic Panel provides researchers with a representative sample of reunited Germany, as well as a longitudinal sample of the western states of Germany since 1984, a longitudinal sample of foreign workers in those states since 1984, a newly begun longitudinal sample of immigrants to those states since 1984, and a longitudinal sample of the eastern states of Germany since 1990. This paper presents an overview of these data and a detailed discussion of the reason for the new immigrant subsample.

1. Introduction

The 1995 version of the German Socio-Economic Panel (GSOEP) is a representative sample of the reunited Federal Republic of Germany in 1995 and provides cross-sectional data on the population of the country as a whole, the population in its western states, foreign workers in those states, recent immigrants to those states, and of the population living in its eastern states. More importantly it is also a longitudinal micro-database of those living in private households in the four samples that make up the full GSOEP. The questionnaires for the western states of Germany were developed and pre-tested in 1983 and the first wave of data was collected in 1984 on this population and on an oversample of guest workers living in the western states of Germany. The 1995 version of the GSOEP contains 12 waves of these data. A sample of immigrants to the western states of Germany since 1984 was added in 1995 to make the sum of the three samples a representative sample of the 1995 population in the western states of Germany in the cross section. A representative sample of the population of German citizens living in the eastern states of Germany was added in 1990 and in the 1995 version of the GSOEP six waves of these data were available. The 1995 version of the GSOEP has been distributed to hundreds of researchers throughout the world.

The GSOEP was funded from 1982 to 1989 by the German National Science Foundation under the auspices of the Special Research Unit entitled "Micro-Analytical Bases for Social Politics." Since then it has been financed by the Federal and State Commission for Educational Planning and Research Support. The German National Science foundation, however, retains administrative control of the

data and its scientific uses. The German Institute for Economic Research in Berlin (DIW) is co-sponsor of the GSOEP.

With funds from the National Institute on Aging — Program Project Nr.: # 1-PO1-AG09743-01, "The Well-Being of the Elderly in a Comparative Context" — the Center for Policy Research at Syracuse University, in cooperation with the German Institute for Economic Research, has prepared an English language version of the Public Use File of the GSOEP. This 95 percent sample of the GSOEP also contains most of the variables contained in the original version. This English language public use version of the GSOEP was created for use by the international research community. Its use, however, is restricted to those who agree to obey the German data protection laws. The DIW and Syracuse University have also collaborated on the creation of the "PSID-GSOEP Equivalent Data File." This comparative data base uses subsets of the United States Panel Study of Income Dynamics and the public use version of the GSOEP. Releases of this data file in 1997 contained 11 years of matched data on income, earnings and various socio-economic variables created from these two data sets and was made available to over 100 researchers throughout the world¹.

2. The Sample

The first wave of GSOEP data, collected in 1984 in the western states of Germany contained 5,921 households. All persons in a household aged 16 and over were surveyed. In

¹ For a more detailed discussion of the PSID-GSOEP Equivalent Data File, see Burkhauser, Butrica, and Daly (1997).

addition to being a representative sample of the population living in these states, the GSOEP was also designed to allow separate analyses of the five largest groups of foreign nationals living in these states. In wave one 1,393 households whose head was from one of these foreign national groups were interviewed. Hence, the GSOEP originally contained two randomly sampled sub-groups:

— *Western States of Germany Sample (Sample A)*: People in private households in the western states of Germany in 1984 where the head of household was not of Turkish, Greek, Yugoslavian, Spanish, or Italian nationality. Other foreign nationals are included in this sample but they made up an insignificant portion of the population in 1984.

— *Foreign Sample (Sample B)*: People in private households in the western states of Germany in 1984 where the head of household was of Turkish, Greek, Yugoslavian, Spanish, or Italian nationality.

The oversampling of foreigners was taken into account in the weighting scheme, to yield a representative sample of the entire population living in the western states of Germany in 1984. In each year since 1984, the GSOEP has attempted to re-interview original sample members unless they left the country or died. Since those aged 16 or older are interviewed in every household, a new generation of panel children is automatically brought into the sample. In addition, new people are surveyed when they join an original panel household.

While Samples A and B of the GSOEP were representative of the population living in the western states of Germany in 1984, they are not fully representative of the new waves of immigrants that have since 1984 come to the western states. Hence, a new sample of immigrants was collected in 1994 and 1995. This sample contains individuals living in 522 households. This new sample is discussed in detail below. The sample is defined as:

— *Immigrant Sample (Sample D)*: People living in private households in the western states of Germany in 1994 or 1995 containing an immigrant who came to the western states of Germany after 1984 (excluding migrants from East to West Germany after reunification, who are covered by Sample C).

The reunification of Germany required another expansion of the GSOEP to make it representative of the whole of Germany. In June 1990 the first wave of a sample of German citizens living in the eastern states of Germany was created. The first wave contained 2,179 households. The sample is defined as:

— *Eastern States of Germany Sample (Sample C)*: People in private households in the eastern states of Germany in June 1990 where the head of household was a German citizen.

3. Survey Instruments

The variables in the GSOEP files come from different survey instruments. In addition to the address protocol

(cover sheet), which the interviewer fills in, the GSOEP also contains household and individual questionnaires. The head of household or the person that knows details about the financial situation in the household answers the household questionnaire. This questionnaire is scheduled to take about 15 minutes to complete. It concentrates on housing quality, income, transfer payments and savings at the household level. For all households with children below age 16 there are questions referring to their kindergarten and school attendance. The information on children is transformed into personal information which is stored in separate "child files." See the Haisken-De New, Büchel, and Wagner, and Spiess papers in this volume for examples of analysis using these data on children.

All respondents aged 16 and over answer the individual questionnaire. This questionnaire is scheduled to take about 35 minutes to complete. It includes approximately 90 questions, including calendars referring to monthly work force participation and various income sources, standard demographic information, objective measures of use of time, earnings, income, benefit payments, etc., and subjective measures like level of satisfaction with various aspects of life, hopes and fears, political involvement, etc. The first time a respondent is interviewed he or she also completes a Biographical Questionnaire, which takes about 30 minutes to complete and provides retrospective information on the person's education, family history, and work history.

All questionnaires are designed either to be administered verbally or to be filled in by the respondents with the assistance of the interviewer. The majority of interviews are done fact-to-face. In exceptional cases people are interviewed by phone or mail. The survey instruments in the Foreign Sample are offered in German or in the respondent's native language.

4. Development of the Samples

The attrition rate has been relatively low in the GSOEP. In 1984, Samples A and B contained 5,921 private households. Each of the 12,245 respondents aged 16 or over in these households was interviewed. Information on the 3,915 children (aged 15 and under) was collected from household heads. The 1995 version of the GSOEP contains 12 waves of Samples A and B. By wave 12, 44 percent of the original 1984 panel members (including children) still continued to have longitudinal records with no missing years.

While attrition has reduced the number of original adult panel members in Samples A and B, the number of individuals interviewed in subsequent cross-sectional samples has been stable due to the entry of panel children who reach age 16 and new members of panel households. There were 8,798 respondents aged 16 and over in 1995 in Samples A and B.

The first wave of Sample C in 1990 contained 2,179 private households with 4,453 respondents aged 16 and

over and 1,591 children. The 1995 version of the GSOEP contains six waves of Sample C. By wave 6, 68 percent of original 1990 panel members (including children) still continued to have longitudinal records with no missing years. There were 3,892 respondents aged 16 and over in 1995 in Sample C.

The 1995 version of GSOEP contains the first full wave of Sample D. It includes 522 private households with 1,078 respondents aged 16 and over and 517 children. Although the 1996 wave is not yet available to the research community, based on the interview counts there will be 1,023 respondents aged 16 and over, plus 469 children in Sample D in that wave.

Users of the 1995 version of the GSOEP have at their disposal a representative sample of reunited Germany, containing information on 6,968 households with 13,768 respondents aged 16 and over and 4,158 children. However, design weights must be used to aggregate the unified data containing all four samples. In addition, cumulative panel mortality is large enough to make it advisable to use weighting factors to cover the combined effect of sample design and panel mortality (see Rendtel 1995 for more details).

Data Protection and Access to Data

Privacy must be guaranteed to insure the long-term cooperation of respondents. In order to make the GSOEP available to the worldwide academic community, including those working in countries whose data protection legislation is less strict than in the Federal Republic of Germany, it was necessary to restrict access to some of the information collected in the GSOEP. The German Institute for Economic Research (DIW), in cooperation with the Center for Policy Research at Syracuse University, has developed such a special version of the GSOEP, which can be used by researchers outside of Germany. In recognition of the international demand for the GSOEP Public Use File, all its documentation is in English. This version, called the "GSOEP Public Use File," is distributed by Syracuse University and is available to any researcher or group of researchers who request the data for academic research and who sign a contract with the DIW to respect the German data protection rules.

As discussed above, to reduce the risk of reidentification, the GSOEP Public Use File omits a few variables contained in the original data. These variables are not critical for most analytical purposes. For example, in Sample B the nationality of a foreigner is aggregated to an indicator that identifies them as not German. In addition, the geographic location of the household and precise asset values of the household are aggregated to broad categories. In addition, only a 95 percent sample of the full GSOEP is made available in the Public Use File². Removing these cases makes it more difficult to identify whether a given individual

is present in the sample, resulting in a smaller risk of reidentification.

5. How and Why Sample D was Created

As discussed above, the major new innovation of the 1995 version of the GSOEP is the inclusion of Sample D. Below we discuss in more detail how and why this sample was created.

Scope and Implementation

Immigration has been changing the demographic composition of Western Europe over the last half century. But political and social changes in Eastern Europe in the last decade have generated a significant new migration to the West, with Germany experiencing by far the largest inflow of immigrants. Münz and Fassmann (1994) report that between 1950 and 1992 68 percent of all east to west migrants to Europe settled in the western states of Germany. In addition to external migration, German reunification caused a tremendous flow of Germans from its eastern to its western states. Between 1988 and 1994 about 1.5 million Germans moved from the eastern to the western states of Germany. In total, about 5.6 million people migrated to, and remained in, the western states of Germany between 1984 and 1994 (cf. Burkhauser et al. 1996; Dunn et al. in this volume).

To capture the consequences of this dramatic change in the German population over the last decade, a sample of these recent immigrants was added to the German Socio-Economic Panel (GSOEP) in 1994 and 1995. This sample will allow researchers to investigate the integration of these immigrants into German life. The next sections provide users of this new data sample with an overview of German immigration policy over the last decade, and describe the characteristics of the immigrant sample.

German Immigration Policy³

In 1950 foreigners made up only about 1 percent of the population in the western states of Germany. This was quite a small share compared to most other European countries. However, that dramatically changed over succeeding decades. Except for the United Kingdom, the foreign population grew substantially in all the countries of Europe. But the foreign population in Germany skyrocketed so that by 1993, foreign residents made up 8.6 percent of the population of reunited Germany.

² One cannot simply take 95 percent of the cases in each wave and use this as a subsample of the panel. This would destroy the longitudinal aspect of the information. The Public Use sample is therefore based on a selection of households from the first wave of the subsample that was interviewed. The data for these households (and all "split-off" arising from them) are contained in all subsequent waves in the Public Use File.

³ For more details see Burkhauser, Kreyenfeld, and Wagner (1996).

Yet even this enormous growth in its foreign population understates the number of new residents that have come into the western states of Germany over the last decade. All "ethnic" Germans who moved from Eastern Europe to the western states of Germany were immediately granted full citizenship and all ethnic Germans who came from the communist controlled eastern states of Germany were already considered German citizens. Neither group is included in the count of foreigners discussed above. When these ethnic Germans are included, the total number of immigrants living in the western states of Germany rises to over 10 million, of which over 5 million arrived between 1984 and 1994 (Schulz 1994)⁴.

In addition to the migration of ethnic Germans to the western states of Germany, the growth in the immigration population was stimulated by permitting temporary "guest workers" to become permanent residents, granting asylum to political refugees, and agreeing to the European Union (EU) protocol of free movement of laborers among EU countries. The impact of each of these forces on Germany immigration is discussed below.

The German Concept of Ethnicity

German citizenship is based on a common cultural, linguistic, and ethnic past. This concept of citizenship is an outgrowth of German history, when "political fragmentation led Germans to think of their nation not as a political or geographical unit, but as a cultural, linguistic, and ethnic one" (Hailbronner 1992). The policy implication of this concept of citizenship is that German citizenship is possible for those who can trace their ancestry to German roots. While non-ethnic persons, even when they are born in Germany, do not automatically receive German citizenship, a *jus sanguinis* (right of blood) allows people of German origin who were born and who live outside the borders of Germany to claim German citizenship.

In practice, however, this right of blood applies almost exclusively to ethnic Germans in Eastern Europe and ethnic Germans who lived under communist rule in the eastern states of Germany. Ethnic Germans from Eastern Europe who resettle in the western states of Germany are called *Aussiedler*⁵. Ethnic Germans from the eastern states of Germany are known as *Übersiedler*⁶. About 53 percent of immigrants to the western states of Germany over the last decade were either *Aussiedler* or *Übersiedler* (Schulz and Seiring 1994).

Ethnic Germans from Eastern Europe (*Aussiedler*)

The German constitution gives certain categories of people of German origin the right to citizenship. This applies mainly to ethnic Germans who lived in communist countries, where they were assumed to suffer from persecution, suppression, or expropriation because of their Ger-

man origin. These ethnic Germans are entitled to live in Germany and are considered citizens from the day they resettle in Germany. The flow of *Aussiedler* into Germany and the countries from which they emigrated varied substantially between 1950 and 1994. The major causes of this erratic migration pattern were changes in political conditions in Eastern Europe.

Between 1984 and 1994 more than two million *Aussiedler* immigrated to Germany and they make up about 2.5 percent of the current population of the reunited Germany (Waffenschmidt 1995). The immigration of *Aussiedler* from Poland and Romania peaked in the years between 1988 and 1991 and has since stopped. Since 1992 *Aussiedler* have come almost exclusively from the former USSR (most especially from Kazakhstan). *Aussiedler* composed 28 percent of all immigrants into the western states of Germany between 1984 and 1994 (Schulz and Seiring 1994). The increase of *Aussiedler* flowing into Germany led to changes in the laws in 1990. The entry procedures for *Aussiedler* were altered to require new immigrants to apply for entry before arrival. In 1992, the influx of *Aussiedler* was limited to a maximum of 220,000 per year (the average number who applied for entry in 1991-1992).

Ethnic Germans from the Eastern States of Germany (*Übersiedler*)

Because the Federal Republic of Germany never recognized the legitimacy of the German Democratic Republic, the German constitution specifies German nationality, not "West" German nationality. From the point of view of the Federal Republic of Germany, those ethnic Germans living in the eastern states of Germany never lost their German citizenship. Consequently they have always been free to resettle in the western states of Germany. *Übersiedler* are defined as ethnic Germans who left the eastern states to resettle in the western states. Such persons were immediately given a German passport and full political and legal rights.

⁴ As will be discussed in detail below, the share of foreigners (non-citizens) in Germany is high because not only for even permanent residents citizenship is difficult to get, but because, unlike many other countries (e.g., the United States, France, etc.) where citizenship is automatically given to anyone born in the country, the children of foreigners born in Germany are not automatically given citizenship.

⁵ The term, *Aussiedler* more precisely refers to ethnic Germans who immigrated between 1953 and 1993. Ethnic Germans who immigrated to Germany after 1993 are called *Spätaussiedler*. Ethnic Germans who immigrated, were expelled, or fled from other countries before 1953 are called *Vertriebene* (expelled). For purposes of this paper, we will call them all *Aussiedler*.

⁶ Following reunification, the legal term for ethnic Germans who moved from the eastern to the western states of Germany changed, but for simplicity we will call all ethnic Germans who lived in the eastern states and moved to the western states of Germany *Übersiedler*.

Emigration from the eastern to the western states of Germany also varied greatly over the past 40 years and was primarily related to the willingness of the communist regime to allow it to occur. The first peak in this type of emigration occurred in the 1950s in the aftermath of the uprisings against the communist government. The construction of the Berlin wall decreased the inflow of *Übersiedler* from more than 230,000 people in 1961 to 21,500 people in 1962. In 1988, with the beginning of the collapse of communism in Hungary, *Übersiedler* fled through Czechoslovakia and Hungary to the western states of Germany. In 1989, with the opening of the borders, the influx of *Übersiedler* increased immediately. In 1989 and 1990 about 800,000 *Übersiedler* migrated to the western states of Germany. This two-year total was more than the 616,000 who migrated during the 27 years of the wall's existence. *Übersiedler* composed about 25 percent of all immigration into the western states of Germany between 1984 and 1994 (Schulz and Seiring 1994).

Foreign Workers of the 1950s and 1960s

In the late 1950s, the Federal Republic of Germany established a worker system to ease its labor market shortages. Treaties with Italy, Turkey, Yugoslavia, and other Mediterranean countries led to huge inflows of foreign laborers. In 1970 almost 2.9 million foreigners, 4.9 percent of the total population, lived in the western states of Germany (Fassmann and Münz 1994). The economic recession of 1973 ended this labor recruitment policy, but foreign workers were not required to leave Germany. Those with work permits were permitted to stay even if their residence permit had expired. After eight years of residence they could apply for an unlimited residence permit, and their families were permitted to join them. Relatively high fertility rates and family migration have steadily increased the relative share of the foreign population in the western states of Germany since 1973. Today these foreign workers and their families have the same economic (but not political) rights as Germans. While they are not entitled to vote, they are eligible for social security benefits (social benefit, unemployment money, pension, etc.).

Granting Asylum

The German constitution of 1949 provides that asylum must be granted to anyone who is persecuted in his country of origin for political reasons. Those granted asylum get the same economic (but not political) rights as Germans. They are, as are *Aussiedler*, entitled to German language courses. In 1992 only about 4 percent of all asylum seekers were granted asylum⁷. Individuals who are fleeing war or social upheaval are not granted asylum. However, because of the Geneva Convention, they cannot be refused entry and a limited level of protection. They are permitted to stay in the country and are only permitted to work on a job that no German or European Union resident would take. The

employment agencies that administer these cases enforce these rules very strictly. About 63 percent of all asylum seekers receive this limited right to stay. They are accommodated in transition camps and receive free board, lodging, and pocket money. The other 33 percent who are denied even limited asylum must return to their country of origin.

Germany has accepted the majority of asylum seekers going to European countries between 1988 and 1993. The demand this large number of asylum seekers puts on resources led to a change in the German asylum law. In 1993, individuals from "safe countries" were no longer allowed to apply for asylum. (Safe countries are defined as signatories to the Geneva Convention, all European Union member states, and a number of other countries). A 1996 decision by the Supreme Court of Germany held this law to be constitutional.

Asylum seekers composed about 14 percent of all immigrants to the western states of Germany between 1984 and 1994.

European Union Membership

The Single European Act of 1993 guaranteed the free movement of people, goods and services among European Union member states. In 1996, members included Belgium, France, Germany, Italy, Luxembourg, the Netherlands, Denmark, Republic of Ireland, the United Kingdom, Greece, Portugal, Spain, Austria, Finland, and Sweden. Citizens of member states are entitled to work in Germany and to receive a residence permit. Residence permits expire three months after they lose or leave a job. They are entitled to unlimited residence after living in Germany for eight years if they are still holding a job.

The Immigrant Sample of GSOEP

When it was first drawn in 1984 the GSOEP was a representative sample of the indigenous population of the western states of Germany. However, between 1984 and 1994 households were added only if a person moved from an original panel household into a new household. Consequently, a representative sample of immigrants arriving after 1984 was not captured in the GSOEP⁸. To overcome this problem, a special immigration sample, Sample D (*Zuwandererstichprobe*), was interviewed in 1994 and 1995. Since 1995 this immigrant sample has been a regular part of the GSOEP.

⁷ Eventually about eight percent of asylum seekers are granted asylum, since once rejected, asylum seekers can take legal action to reverse the decision.

⁸ The probability that the GSOEP contains immigrants who arrived after 1984 is not zero. In wave 12 Frick and Wagner (1996) found 34 new immigrants in Sample A, 208 new immigrants in Sample B, and 204 new immigrants in Sample C.

As was the case with Sample B, which contains foreign workers, the most recent immigrants were oversampled in order to allow researchers sufficient sample size to characterize them. Sample D includes households that contain at least one member who immigrated to the western states of Germany between 1984 and 1994. Interviews were conducted solely in the western states of Germany because few people immigrated to the eastern states of Germany over this period. People living in public institutions (nursing homes, mental institutions, etc.) or asylum camps are not included in the first wave.

Sample Selection

Sample D was planned to be a random sample of 500 recent immigrant households. Survey planning was hindered by the absence of an official government register from which to draw addresses of immigrants. There are central registers for non-German residents, but none for *Aussiedler* or *Übersiedler* who are recognized as Germans upon entering the country. Moreover, there is no official census of the population of immigrants. To locate households of immigrants the polling organization *Infratest Sozialforschung* carried out address screenings in 1992 and 1994. Households gathered in the first address screening are labeled D1, and those gathered in the second address screening are labeled D2. This two-stage procedure was done because there were not sufficient funds to do it in one step. It turned out that this two-step-procedure was useful since the experience gained from the first round significantly improved overall sampling procedures.

First Address Screening (D1)

In the first address screening, two address selection methods, *Infratest-Bus* and *Infratest-Scope*, were applied, as was a "snowball" system to gather further addresses. *Infratest-Bus* used a random-route-walk method to find household addresses within voting districts. The sample plan is based on a design by *Arbeitskreis Deutscher Marktforschungsinstitute (ADM)*. *Infratest-Scope* used a random phone call method to find household addresses based on a random dialing algorithm of telephone numbers. Both methods can cause selection problems to the degree that voter registration or telephone ownership are not themselves random in the population.

The left panel of Table 1 provides information about how the initial D1 sample was formed. Using these two methods, households were selected from an *Infratest Sozialforschung* run market research interview in 1992. At the end of the interview the household member interviewed was asked the following question: "Is there currently an adult in your household who lived in a foreign country or in the eastern states of Germany in 1984?"

There were 28,776,000 non-institutionalized households in the western states of Germany in 1992. Of these, 7,956 or .028 percent were asked this question via random-route-

Table 1
Random Sample Frames and Identification of the Immigrant Subsamples

Households	Sample D1 (1992) ¹⁾		Sample D2 (1994) ²⁾	
	Number	Percentage	Number	Percentage
Total	28,776,000	100	29,562,000	100
Screened	7,956	.00028	11,000	.00037
Identified as New Immigrant	330	4.1	725	6.6
Willing to be interviewed	195	59.1	548	75.6
<i>Aussiedler</i>	86	—	180	—
<i>Übersiedler</i>	83	—	146	—
All others	26 ²⁾	—	198	—
Approached for interviews	195	—	400	—
<i>Aussiedler</i>	86	—	125	—
<i>Übersiedler</i>	83	—	77	—
All others	26 ²⁾	—	198	—

1) This includes the *Infratest-Bus* and *Infratest-Scope* Sample only. In 1994, 141 additional households were identified as new immigrant households using a snowball method. Of these, 98 households were willing to be interviewed and were interviewed. — 2) Includes nine households whose original country could not be identified.

walk or telephone interviews. Approximately 4 percent (330) answered "yes," of whom 195 or 59 percent agreed to be interviewed at a later date. Of these the great majority were either *Aussiedler* or *Übersiedler*. In 1994, 138 of these households were successfully interviewed and make up the random interview component of Sample D1 (see column 1 of table 2). To reduce costs and to facilitate address gathering, a snowball system was then used to increase sample size. This was done by asking the 138 immigrant households if they could provide the addresses of any

Table 2
Size and Distribution of 1994 and 1995 Surveys of Immigrant Sample

	D1 Sample			D2 Sample	Total Immigrant Sample
	Random Sample	Snowball Sample	Total Sample ¹⁾	Random Sample	Total Sample ¹⁾
Wave 1 (1994)					
<i>Aussiedler</i>	67	65	132	—	132
<i>Übersiedler</i>	64	30	94	—	94
All Others	7	3	10	—	10
Total	138	98	236	—	236
Wave 2 (1995)					
<i>Aussiedler</i> ²⁾	60	52	112	104	216
<i>Übersiedler</i>	61	31	92	59	151
All Others ²⁾	11	12	23	132	155
Total ²⁾	132	95	227	295	522

1) Includes "snowball" sample population. — 2) The sample in 1995 includes new households created by splits of the original population.

immigrant friends. This method yielded 141 additional households and 98 of whom (see column 2 of Table 2) were also interviewed in 1994. Based on these two methods of address gathering, a total of 236 households (see column 3 of Table 2) were interviewed in Sample D1 in 1994.

Voter registration lists were used to draw the initial *Infratest-Bus* sample points. Because *Aussiedler* and *Übersiedler* can vote immediately upon establishing residence in Germany while other new immigrants cannot, sample points with *Aussiedler* and *Übersiedler* are over-sampled using this method (see Schulz et al. 1993 for a full discussion). In addition, the snowball system also over-represented *Aussiedler*, since they tend to immigrate in networks (Bauer and Zimmermann 1995), while *Übersiedler*, asylum seekers, and people from European Union-member states are less likely to rely on migration networks. These differences in sampling probabilities are taken into account by a reweighting process (see below).

Second Address Screening (D2)

In 1994 *Infratest* started another address screening. Based on the experience of D1, only the *Infratest-Bus* random-route-walk method was applied. As can also be seen in Table 1, of the 11,000 households initially interviewed, 725 immigrant households were identified. From these 725 households, 548 agreed to further interviews. In order to correct for their overrepresentation in D1, 148 *Aussiedler* and *Übersiedler* households were randomly excluded. From the remaining 400 addresses, 295 immigrant households agreed to further interviews (see column 4 of Table 2). These 295 immigrant households make up sample D2. (For more details on issues raised by the address screening procedures in Sample D, see Schupp and Wagner 1995).

Sample Sizes

The first survey of the immigrant sample was fielded in 1994. This sample includes 138 households that come from the 195 households screened via *Infratest-Bus* and *Infratest-Scope* who agreed to be interviewed in 1992, as well as the 98 households identified in 1994 via the snowball method who agreed to be interviewed. As can be seen in column 3 of Table 2, when these two groups are combined the total sample size in D1 is 236 households.

Because Sample D2 was not interviewed until 1995, the 1994 survey of immigrants consists solely of the D1 Sample. This is not a representative sample of the recent immigrant population. Since no sample weights were assigned to this 1994 sub-sample, it can not be used to describe the entire population of recent immigrants, except in conjunction with the 1995 data, which is weighted.

The distribution of immigrant subgroups for Sample D2 is also reported in Table 2. Of the remaining 400 households in 1994, 295 were successfully interviewed again in 1995,

the majority of whom were *Aussiedler* and *Übersiedler*. When households from D1 and D2 are combined, wave 2 of the immigrant sample consists of 522 households. When the snowball households are excluded and only the random sample is used, the sample falls to 427 households.

Case Weights

Because some of the households in the immigrant sample could also be captured in the other GSOEP-Samples it was necessary to create two different sets of weights to integrate the immigration sample into the GSOEP design. The first set is to be used with the full GSOEP and the second is to be used when working with Sample D alone. When the full GSOEP file is used, the conventional cross-section personal weight (wave)PHRF and household weight (wave)HHRF are appropriate. When only Sample D is used, the weights whose names are extended by "D" are appropriate, i.e., the cross-section individual weight (wave)PHRFD and household weight (wave)HHRFD. This second set of weights assigns a nonzero weight to any household group in sample D that is also represented in the other GSOEP-Samples. Table 3 provides a more precise verbal definition of these weights.

Households in Sample D that were gathered through the snowball method are excluded from the weighting procedure since it is difficult to control for their probabilities

Table 3
Definition of Sample Weights in Wave 12¹⁾

LPHRF	Person level weight for cross-sectional analysis for wave 12 (1995), if all samples (A, B, C, and D) are used. A zero household weight is assigned to all persons in households in Sample D who live in a household that contains both an immigrant and a non-immigrant (mixed-households).
LHHRF	Household level weight for cross-sectional analysis for wave 12 (1995), if all samples (A, B, C, and D) are used. A zero household weight is assigned to all persons in households in Sample D who live in a household that contains both an immigrant and a non-immigrant (mixed-households).
LPHRFD	Person level weight for cross-sectional analysis for Wave 1 (1995), if only Sample D is used.
LHHRFD	Household level weight for cross-sectional analysis for Wave 1 (1995), if only Sample D is used.

¹⁾ For those who want to compare Wave 12 with previous waves, which contain only Sample A and Sample B people living in the western states of Germany, the LPHRF and LHHRF weights must be multiplied by 1.053 to reflect the loss of immigrant population that occurs when Sample D people are not included. For a more detailed discussion see Rendtel, Pannenberg and Daschke (1997).

(Rendtel, Pannenberg and Daschke 1997). They are assigned a zero weight; because the *Übersiedler* who moved to West Germany after unification are sampled on a selective basis only they are assigned a zero value too (but these households can be used for multivariate regression analysis). Because only 11 recent immigrant households in Sample D1 were neither *Aussiedler* or *Übersiedler*, the weights for the "other new immigrant" group were constructed based on sample D2 alone.

Because there is no official government count of immigrants, the German Institute for Economic Research made its own estimation on the basis of several data sources to externally validate the sample (Schulz 1994). Table 4 compares these externally obtained values with the sets of values derived from the 1995 wave of Sample D. The total estimated new immigrant population and the share of that population made up of *Aussiedler*, *Übersiedler*, and all other immigrants are estimated using the unweighted sample, using the assigned design weights developed above. When the unweighted sample values are compared with the externally obtained values, it is found that Sample D contains, as expected, an oversample of *Aussiedler* and an undersample of all other new immigrants.

When the original design weights are used to correct this problem, the results more closely resemble the externally estimated household population. But even using these weights, the number of recent immigrants is understated and the distribution of subgroups is not exact. In order to more closely match the externally obtained total population of recent immigrants, it was decided to fit the original design weights to the external distribution given by Schulz

Table 4
Comparison of the Size and Subgroup Shares of the 1995 Immigrant Sample, Unweighted, Weighted and Compared to External Information

	Unweighted ¹⁾	Design Weight Size and Share Values	
		Original	External Values
Total Household	378	1,639,000	1,980,000
<i>Aussiedler</i>	43.4	35.2	29.3
<i>Übersiedler</i>	21.7	21.1	18.7
All other new immigrants	34.9	43.7	52.0
Total	100	100	100

¹⁾ The snowball sample households and households of *Übersiedler* who moved after unification were given a weight of zero and effectively ignored in these calculations. The same is true of the 11 new immigrant households who were neither *Aussiedler* or *Übersiedler* in the 1995 D1 sample.

Source: Rendtel, Pannenberg and Daschke (1997)

(1994) with respect to the different subgroups. In fact, there are only minor changes for the subgroups *Aussiedler* and *Übersiedler* (0,5 %; 6 %), but the weighting factors for the group of all other immigrants are increased by 1.44. (For a more detailed discussion of how the weights were developed, see Rendtel, Pannenberg and Daschke 1996.)

Table 5
Sample D — Biographical File Variables

Immigration History

German nationality?
Wish to become German citizen?
Attitude towards dual citizenship
Where lived in 1984?
Year of arrival
Country of birth
Permanently living in Germany?
Status at arrival (e.g., *Aussiedler*, *Übersiedler*)
Limited or unlimited residence permit?
Lived in transition camp in the beginning?
Time in transition camp

Family Network

Came to Germany as family member of someone already in Germany?
Before arrival: any contact with a family member already in Germany
Moved to the place friends/relatives are living?
Kind of help received from relatives/friends
Reasons to resettle
Expectations of Germany
Family members still in country of origin?
Wish them to follow here?
Plan to go back someday?
Circle of acquaintances: Nationality/sex/family/relation

German Proficiency

German Language course?
Language spoken in Germany
Ability to speak German
Ability to write German
Ability to speak the native language
Ability to write the native language

Discrimination

Feel discriminated against?
While looking for a flat / job / at offices / while shopping?

Schooling

Father's schooling
Mother's schooling
Attended school in Germany
School degree
Years of schooling in another country
Type of school in another country
Cont. educational school degree (another country) recognized in Germany

Variables

In addition to the standard questions asked of all members of the GSOEP, respondents in Sample D were asked a series of questions in their Biographical Questionnaires that probed specifically at issues related to their immigration experiences. Table 5 provides a selection of these questions. These variables were not made available on the 1995 version of the GSOEP but can be obtained from the German Institute for Economic Research.

6. Summary

The 1995 version of the German Socio-Economic Panel Study provides a representative sample of reunited Germany in 1995, as well as a longitudinal record of the population of its western states back to 1984 and its eastern states back to 1990. The dramatic increase in the immigrant population in the western states of Germany between 1984 and 1994 made it necessary to add a sample of recent immigrants to the GSOEP western state sample to make it representative in the cross section of the population in 1995. The addition of Sample D achieves this goal. And when Sample D is combined with the other three GSOEP sample populations, the entire GSOEP sample is representative of all people living in private households in Germany in 1995. Furthermore, researchers using Sample D alone can perform structural and longitudinal analysis specifically on the population of recent immigrants to the western states of Germany. However, because *Aussiedler* and *Übersiedler* are oversampled in Sample D, weights are necessary for descriptive analysis using these data.

This paper provides readers with a brief description of the GSOEP, the Public Use version of the GSOEP and the PSID-GSOEP Equivalent Data File as well as a more detailed discussion of Sample D, the reasons for its creation, the way it was created, and some of the variables it contains. While these data are relatively new, the special reference section at the end of this volume shows that the GSOEP has already had an effect on international research in the social sciences. This special section contains a selection of English language research papers using the GSOEP.

How to Obtain the Data

German users can receive the full sample of the GSOEP from the German Institute for Economic Research at Berlin. All users outside of Germany must get the Syracuse University Public Use Version of the GSOEP through Syracuse University. The dissemination of the Syracuse University GSOEP Public Use File is a two-step process. First it is necessary to sign a contract with the original producer of the data, the German Institute for Economic Research. After the contract is signed, the data is disseminated through Syracuse University. The charge for processing orders for the GSOEP data is 200 DM. The charge for processing the Syracuse University Public Use Version of the GSOEP is \$125. The public use version of the data comes with extensive documentation (written in English) on a CD-ROM. By means of a setup program the data can be copied to any platform in SPSS, SAS, or ASCII-format.

The Syracuse University PSID-GSOEP Equivalent Data File is automatically provided free of charge on the same CD-ROM to users who request the Public Use Version of the GSOEP.

For more information on how to order these data sets, please contact either:

Deutsches Institut für Wirtschaftsforschung
SOEP Secretary
Koenigin-Luise-Strasse 5
D-14191 Berlin

Tel: +49 30 897 89 283
Fax: +49 30 897 89 200
Internet: SOEPMail@DIW-BERLIN.DE

or:

Mr. Detlef Jurkat
Center for Policy Research
Syracuse University
426 Eggers Hall
Syracuse, NY 1322-1090

Tel. +1-315-443-9442
Fax: +1-315-443-1081
E-mail: dwjurkat@syr.edu

References

- Bauer, Thomas and Klaus F. Zimmermann.* 1995. "Integrating the East: The Labor Market Effects of Immigration," *Münchener Wirtschaftswissenschaftliche Beiträge*, 95 (15). Munich: mimeo.
- Burkhauser, Richard V., Michaela Kreyenfeld, and Gert G. Wagner.* 1997. "The Immigrant Sample of the German Socio-Economic Panel," Cross-National Studies in Aging Program Project Paper No. 29, All-University Gerontology Center, The Maxwell School. Syracuse, NY: Syracuse University, forthcoming.
- Burkhauser, Richard V., Barbara A. Butrica, and Mary C. Daly.* 1997. "The Syracuse University PSID-GSOEP Equivalent Data File: A Product of Cross-National Research," Cross-National Studies in Aging Program Project Paper No. 25 [revised], All-University Gerontology Center, The Maxwell School. Syracuse, NY: Syracuse University.
- Dunn, Thomas A., Michaela Kreyenfeld, and Mary E. Lovely.* "Communist Human Capital in a Capitalist Labor Market: The Experience of East German and Ethnic German Immigrants to West Germany," in this volume.
- Frick, Joachim and Gert G. Wagner.* 1996. „Zur Sozio-ökonomischen Lage von Zuwanderern in Westdeutschland," DIW Diskussionspapier. Berlin: mimeo.
- Hailbronner, Kay.* 1992. "Citizenship and Nationhood in Germany." In W. Brubaker (ed.), *The Politics of Citizenship in France and in Germany*. Cambridge: University Press, pp. 67-79.
- Haisken-De New, John P., Felix Büchel, and Gert G. Wagner.* "Assimilation and Other Determinants of School Attainment in Germany: Do Immigrant Children Perform as Well as Germans?" in this volume.
- Münz, Rainer and Heinz Fassmann.* 1994. „Geschichte und Gegenwart europäischer Ost-West Wanderungen." In Münz, Korte, and Wagner (eds.) *Internationale Wanderungen*. Demographie Aktuell, Nr.5. Berlin.
- Rendtel, Ulrich.* 1995. „Panelausfälle und Panelrepräsentativität," Sozio-ökonomische Daten und Analysen für die Bundesrepublik Deutschland, 6. Frankfurt-Main: Campus.
- Rendtel, Ulrich, Markus Pannenberg and Stefan Daschke.* 1997. „Die Gewichtung der Zuwanderer-Stichprobe des Sozio-oekonomischen Panels (SOEP)." *Vierteljahrshefte zur Wirtschaftsforschung*, No. 2 (in print).
- Schulz, Erika.* 1994. „Zuwanderungen der letzten zehn Jahre und Abschätzung des Bestands zum 31.12.93," DIW Diskussionspapier. Berlin: mimeo.
- Schulz, Erika and Kirsten Seiring.* 1994. „Integration Deutscher Zuwanderer in den Westdeutschen Arbeitsmarkt," *Wochenbericht des DIW*, Nr. 35, pp. 609-617.
- Schulz, Erika, Ulrich Rendtel, Jürgen Schupp, and Gert Wagner.* 1993. „Das Zuwanderer-Problem in Wiederholungsbefragungen am Beispiel des Sozio-oekonomischen Panels," Berlin: mimeo.
- Schupp, Jürgen and Gert Wagner.* 1995. „Die Zuwandererstichprobe des Sozio-oekonomischen Panels (SOEP)," *Vierteljahrshefte zur Wirtschaftsforschung*, No. 1, pp. 16-25.
- Spiess, C. Katharina.* "American and German Mothers' Child Care Choice: Does Policy Matter?" in this volume.
- Waffenschmidt, Horst.* 1995. „Info-Dienst Deutsche Aussiedler, Zahlen, Daten, Fakten." Bundesministerium des Inneren, Bonn.