

SOEPpapers

on Multidisciplinary Panel Data Research

163

Hartmut Esser

**The People of "Linden Tree Street"
and the German Socio-Economic Panel (SOEP)**

Berlin, March 2009

SOEPPapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPPapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPPapers are available at
<http://www.diw.de/soeppapers>

Editors:

Georg **Meran** (Dean DIW Graduate Center)

Gert G. **Wagner** (Social Sciences)

Joachim R. **Frick** (Empirical Economics)

Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics)

Christoph **Breuer** (Sport Science, DIW Research Professor)

Anita I. **Drever** (Geography)

Elke **Holst** (Gender Studies)

Frieder R. **Lang** (Psychology, DIW Research Professor)

Jörg-Peter **Schräpler** (Survey Methodology)

C. Katharina **Spieß** (Educational Science)

Martin **Spieß** (Survey Methodology)

Alan S. **Zuckerman** (Political Science, DIW Research Professor)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | urahmann@diw.de

The People of “Linden Tree Street” and the German Socio-Economic Panel (SOEP)

Hartmut Esser

University of Mannheim, Faculty of Social Sciences

Lately I have been receiving numerous requests for contributions to commemorative publications, farewell speeches, and anniversary lectures. I cannot respond to all of these requests, especially since not all of them are as inspiring as the one today: the 25th (in words: the twenty-fifth!) wave of the SOEP! This is indeed an anniversary that should be duly celebrated and honored. The German soap opera “Lindenstrasse” (Linden Tree Street)¹ that was first aired in December 1985 is almost as old as the SOEP. Who could have foreseen in 1984 that there would be so many waves of the SOEP and so many episodes of “Lindenstrasse” following that first one? In an article well worth reading, Hans-Jürgen Krupp (2008) has described the genesis of SOEP, revealing the diverse reasons for initiating this project—some less obvious than others — and describing the resistance it faced on both justified and unjustified grounds, and the great courage and farsightedness of those who championed this cause. It was they who have made SOEP the great success it has become, which nobody could have predicted at that time.

While working on this article² and also reading the one by Hans-Jürgen Krupp, I came to realize that I have had the chance to witness a remarkable moment in the history of socioeconomic research, and that the past 25 waves of SOEP are a

¹ Lindenstraße (“Linden Tree Street”) is considered to be the first “real” German soap opera, and takes its name from the basswood trees lining many German streets. It strives to be as authentic as possible in representing German life, suggesting that the people living on “Lindenstrasse” are “real” and that their hopes and dreams resemble those of the average German. It was also the first German TV series to address serious and controversial issues such as poverty (a major topic of SOEP), AIDS, domestic violence, racism, and homosexuality.

² I am grateful to Sabine Kray, Debbie Bowen and Betty Haire Weyerer for doing a wonderful job in translating my German manuscript into English.

significant milestone in my personal vita. In the following, I will address four different aspects. First, I will describe some of my impressions while participating in the establishment of the SOEP and its first waves. Second, I will draw attention to the SOEP's role in current public debates (such as the one on migration issues). Third, I will describe my experiences as a SOEP user, and finally, I will elaborate on the great significance of the SOEP for the integration of the social sciences and their empirical methodologies.

The first aspect is perforce also (already) somewhat of a review, an often awkward matter (and one that is to be pardoned, because the request for this contribution demanded precisely that).

The SOEP and I

My first contact with what would later become the SOEP took place in 1981, when I was asked by ZUMA if I would consider heading a team that was to develop a large-scale panel study. Mannheim (where ZUMA was located) was the place to be at that time (as it still is today) for someone as interested as I was in the empirical social sciences. But in the end it all worked out differently: the move to ZUMA did not take place in 1981 (and my later work as Managing Director at ZUMA and as a full professor in the Faculty for Social Sciences was filled with other activities and other interests). Nevertheless, SOEP continued to take shape. In these early stages, my role suddenly turned out to have less to do with developing the methodology and more with shaping the survey's content, namely that of evaluating suggestions for the inclusion of migrants in the SOEP ("ethnic boost"). At the time I was making these decisions, I was not yet aware of the great impact they would later have. We arrived at a disproportional sample of the most important immigrant groups (labor immigrants (*Gastarbeiter*) and later Ethnic Germans (*Aussiedler*)) and developed a number of indicators of integration and assimilation that have since been expanded and improved upon and are used widely today. Most scholars now are probably unaware of how far-sighted, but then again also how random, these decisions with regard to immigrant samples were. For the founders of the panel, it was obvious that (West) Germany was an immigration country even before the political and public debates over the issue began to gain momentum. It was also not terribly hard for them to make decisions relating to content once they had decided to commit themselves fully

to this ambitious new project. Yet, once the study was underway, changes and innovations became somewhat more difficult than they had been in the starting phase. That was also my main experience while serving on the SOEP Advisory Council. The struggle to reach compromises between the SOEP team, individual users, and the members of the advisory council was constant, inspiring, and very fruitful. Especially in the early years, the SOEP team's main objective with regard to their project was continuity, thus they were careful about introducing "wild" ideas into the program. As the years passed, the SOEP team became more willing to try out new, unusual ideas, and to combine and connect different topics that had not previously been linked. Today the SOEP is one of the most innovative institutions around. Its list of new and interdisciplinary topics is impressive and proves that the SOEP is not one of those institutions in the social sciences that will die a slow, bureaucratic death.

The results of these enterprising ideas were not always predictable, given the complex nature of the topics and issues brought together within the SOEP. Yet the results (often not intended as such) are impressive and useful, especially in the area of immigration issues. The SOEP data afford an objective, scientific foundation for highly politicized and emotionally charged debates — even though the transmission of these data into the public realm is usually another, more difficult, story.

The SOEP not only offers data concerning the objective and subjective well-being of migrants, but also includes information on language, culture, friendships, networks, and identity. The SOEP methods allow for the respective inclusion of mixed characteristics and orientations, i.e., "multiculturalism", such as bilingualism, ethnically hybrid behavior, mixed friendships, and multiple identities. Recent discussions about multiculturalism have been emotionally charged and ideologically driven, and have raised manifold questions: Is multiculturalism really better for the labor market than assimilation? Are ethnic schools a solution to the problem of ethnic stratification? Is there such a thing as unforced assimilation? Is a religious and ethnic withdrawal among second and third-generation immigrants really under way, as scholars in cultural studies like to argue and most probably would like to see occur? According to these scholars, assimilation in terms of language, culture, friendships, and identities is no longer "in," and young people are tending instead toward what

scholars call re-ethnicization. In most cases, this is nothing but an assumption; the factual evidence generally being derived from case studies, or pure theory and “good intentions.” Systematic examinations of such theories are rare, if at all existent, due to the lack of data on the conditions and development of interethnic relationships across several generations, which would facilitate a systematic examination of such hypotheses. The SOEP data allow for this kind of detailed and sound analysis, and the results are quite surprising: contrary to one’s expectations, the “classical” theories of intergenerational integration still hold true. According to the SOEP data, assimilation in language, cultural patterns, networks, and identities happens gradually from generation to generation. Some ethnic groups are slower to assimilate than others, a prominent example being the Turks, but they too are gradually becoming more and more assimilated.

The reasons for these differences are again those found earlier on: education, age on arrival, length of stay, (de)segregation, interethnic contacts, and other factors. In language assimilation, for example, “multiculturalism” plays a relatively minor role. While this sounds unbelievable to most people, such results become apparent only after a longer period of examination, after looking at several waves. The usual cross-sectional analyses show the situation at a given point—for example, an individual living in a segregated community or taking a trip to his or her home country—but they do not say anything about the duration or development of that situation. Only *permanent* segregation or a *longer* stay abroad will provoke ethnic segmentation, leading, for example, to a stagnation of skills in the second language. With the SOEP such mistakes can be avoided in future – even despite the fact that many researchers continue to regard the SOEP as a giant cross-sectional dataset and thereby, even with this potential, unnecessarily draw the wrong conclusions.

This brings us to the third aspect, the use of the SOEP data. Data management, data distribution, and the proper usage of the panel data as such were the most important challenges from the start and they still are today. Many things have changed but data preparation, data evaluation, and statistical modeling still are very difficult for many data users. At some universities, support staff members who are familiar with the data and with the statistical methodology are more influential than their supervisors. Yet there is no other (or better) solution to the dilemma that handling the data is

difficult, and it has to be said that the one of the most impressive achievements of the SOEP team over time has been to make that easier — despite the dramatic enlargement of the data and its increasingly complicated structure. In the starting phase of the SOEP (once again in reference here to “history”) and while playing various roles simultaneously (Chairman of the SOEP Advisory Council, Member of the GESIS Curatorium, Director of the Institute for Applied Social Research in Cologne), in hindsight I had at the time the somewhat naïve and superfluous, ultimately wrong, idea that the SOEP team could and should not be burdened with data management or data distribution, and that these activities should constitute a genuine task for the GESIS infrastructure, specifically the Central Archive for Empirical Social Research in Cologne. *Both* SOEP and GESIS disagreed: the members of the Central Archive did not want to take on these new challenges, while the SOEP team insisted on doing so. And that is how the current structure evolved: all important tasks are now carried out by one and the same institution that combines methodological and content-related knowledge. Users are highly welcome and treated as customers. The SOEP benefits directly from these “customers” and thus as long as they are able to define what it is that they need, the staff always strives to give them whatever they need, and in the needed form, within the constraints of what is possible. While this has always been my experience, it perhaps cannot be completely generalized. Due to a variety of circumstances, I did not have an opportunity during the early days of institutionally advising the SOEP to use the data intensively myself, and the numerous invitations to journey to Berlin for a SOEP-analysis remained unheeded for many years.

Later, however, I did have the opportunity to work closely with the SOEP team on two occasions: a research stay at the WZB during the fall term 2004/5 and a several-day visit to the DIW in February 2007. Both of these projects were related to the question of integration of immigrants, the results of which I have cited above. In 2007, I discussed questions with the SOEP related to context data known as Geocodes, which at the time (because of the detailed geographical nature of the Microm data) were to be used only on site at the DIW (Microm data). And at both of the above occasions the problem was the same: how to ensure uncomplicated access to the data sets needed and how to eliminate technical errors. Not only did the SOEP team

try to help; they also got involved, made important suggestions, and engaged in very inspiring discussions of the issue with me.³

In the end, the context data on the integration of immigrants did not really lead anywhere in terms of “explained variances”, but that is another story, and one I would have enjoyed telling. The underlying reason soon became apparent: the available context data oftentimes do not correspond to the issues and problems at hand. Also in some cases the context areas are too large or the indicators too indirect to produce productive results. In the case of very small-scale geographical data, such as block-level information on foreigners, it turned out that the available indicator was not capable of representing the variable of interest. But again, this was precisely the problem: until then there were simply no other data available. Today, the SOEP team is undoubtedly hard at work further improving this data situation.

And finally the fourth aspect. It has to do with a completely different matter, another positive development brought about by the SOEP and its initiators: the convergence of the empirical social sciences. As can already be gathered from its name, the SOEP has always been interdisciplinary. Of course sometimes the “S” for “Socio-,” as in social sciences, dominated the SOEP’s activities, and other times it was the “OE,” the economic side of the panel project. The inclusion of psychology at a later stage was not welcomed by all, since, as is in a sense inherent to psychological constructs, for some it seemed to lack sound theoretical and structural reasoning, if not to say that it seemed to be somewhat ad hoc.

Meanwhile the program has expanded its scope much further, now including not only sociology, economics, and psychology, but also demographic, biological, and neurophysiological mechanisms, and elements of empirical game theory. A particularly innovative addition in this context was the theme of limited rationality—the norms, emotions, and cultural influences moderating the effects of (economic) stimuli. What is noteworthy here is that this aspect is based on a full-fledged theoretical conception that offers explicit hypotheses. This prevents social scientists from doing the usual: putting all variables the SOEP offers into the regression equations at once.

³ In the context of these very positive experiences I would like to mention some names. First of all Juergen Schupp for the overall support, then Martin Kroh, who was very helpful during both of the above-mentioned projects, and finally Gundi Knies and Niels Witte who helped me greatly with the context data.

Usually, this leads to terrible confusion, especially when yet another scientist working on the same problem arrives at yet a different result—simply because he/she has used data from different waves, or even the whole data set, to answer that one specific question. And it is exactly here that in addition to such positive “convergences”, other, dubious ones can be found. There are increasing signs that the “imperialism” of some economists extends even beyond simply appropriating classical sociological concepts little by little for their own ends. In so doing they often appear to give even less thought to the unregulated use of variables in their “empirical” pursuits in general than empirical sociology ever has. This effect is exacerbated by the somewhat cumbersome division of labor between sociologists and economists, an exaggerated, but not completely false or unfair, description of which follows. While the sociologists used to neglect the theoretical mechanisms because they were too busy defining the terms, creating the phenomenological description, and conducting frustrating and time-consuming *primary* surveys, the economists devoted themselves to more pleasant tasks: the theoretical modeling and the *secondary* analysis, with all its econometric nuances. The data for their analyses are derived, of course, from the aforementioned *primary* surveys conducted by the sociologists. One could even say that the economists really are free-riders, since they are profiting from the results without any intention to ever undertake such frustrating jobs themselves.

While many sociologists have since learned from their old mistakes and started to make reference to theoretical work as a basis for their empirical research, some economists who have only recently discovered the SOEP data are repeating the same mistake of relying exclusively on the empirical data. This leads to an accumulation of more or less related facts without generating any productive results. The SOEP is not responsible for that.

Concluding Remarks

On “Lindenstrasse,” just as in the SOEP, people have come and gone over the years, while others have stayed on. Also, the topics dealt with on “Lindenstrasse” have

changed over time, as have those dealt with by the SOEP. Yet “Lindenstrasse”, as far as one can judge, has not improved over the years; the protagonists are not more handsome or more beautiful, and the general interest in the series is stagnant. The SOEP, on the other hand, has become better and better with each successive year, as it has grown bigger, more diverse, and more inspiring for researchers. The appreciation for the SOEP in the research community is great, and possibly the greatest compliment of all is that several other institutions are now in the process of creating their own panel studies modeled directly on the SOEP.

The SOEP will continue to provide an outstanding model when it comes to organization, scientific excellence, and above all—its “get up and go” mentality. The aforementioned newcomers are the Relationship and Family Panel (PAIRFAM), which has already entered its starting phase, and the National Education Panel (NEPS), which is soon to be launched. With these two new German panels flanking the SOEP, a truly unique combination of detailed observations of these key realms of society results. I sincerely hope that these two projects will soon be ready to take off, and that they will be just as successful as the SOEP.

Reference

Krupp, Hans-Juergen (2008), The German Socio-Economic Panel: How It All Began. SOEPpaper No. 75. Berlin.