

# Readme Document

on the data used for the paper

## The effect of rapid structural change on workers

by Eva Weigt

### 1. The Socio-economic Panel

The data set used for the analyses of this paper is generated from the Socio-Economic Panel (SOEP). The SOEP is a household panel survey conducted on yearly basis since 1984 in Western Germany by the DIW. In the SOEP, all members of a household aged at least 16 years are included. The data contain socio-demographic characteristics of the respondents as well as, inter alia, detailed information on labour market biographies (including earnings and wages), political orientations and household characteristics. ). For a description of the SOEP data see Wagner et al. (2007).

The first SOEP wave on East German territory was conducted in June 1990 under the title “GDR sample of the German Socio-Economic Panel” and comprised interviews of 4,313 individuals living in 2,130 households (Schupp and Wagner, 1990). In particular, I used the cross-sectional data version 30, SOEP, 2014, doi:10.5684/soep.v30 and restricted the data to first-time respondents of the first wave conducted on Eastern German territory in 1990. It should be noted that the SOEP micro data are not publicly available, but scientific researchers can get access for free. Please visit the website of the SOEP to learn more on how to get access to the SOEP data:

[https://www.diw.de/en/diw\\_02.c.222829.en/access\\_and\\_ordering.html](https://www.diw.de/en/diw_02.c.222829.en/access_and_ordering.html)

### 2. How to work with the provided dofiles

To prepare and analyse the cross-sectional original SOEP data provided by the DIW, I used the statistics software Stata. The respective dofiles used for the preparation and analyses of the data can be found in the data archive.

I used the following five different dofiles to work with the data:

- Weigt01\_Pull\_and\_merge\_SOEP\_data

- The code of this dofile is basically provided by the SOEP Info Website to pull the original single variables from the original cross-sectional data files and merge them into one big data set.
- Weigt02\_Naming\_and\_recoding\_cross\_sectional\_data
  - The second dofile contains the most basic preparation of the original cross-sectional data by renaming and recoding all used cross-sectional variables. Please find below a full description of the used variables.
- Weigt03\_Reshaping\_data\_into\_panel
  - In this dofile the data are reshaped from cross-sectional into panel format. Furthermore, for some variables time-invariant information for the year 1990 is generated to control for these characteristics in the subsequent empirical analyses.
- Weigt04\_Descriptive\_statistics
  - The fourth dofile contains the code for the descriptive analyses shown in the paper.
- Weigt05\_Regression\_analyses
  - The fifth and last dofiles contains the code for the regression analyses shown in the paper.

### **3. Description of variables**

yobirth = "Year of birth"

female = "Sex"

0 = "Male"

1 = "Female"

Note that the following variables are created from the original cross-sectional files provided by the SOEP. They are included separately as single yearly variables in each year. Thus, in the cross-section each variable includes a suffix for the respective calendar year (1990 to 1999). These suffixes are not shown here for the sake of clarity and as they are removed after the reshaping of the data into a panel data set (see dofile Weigt03\_Reshaping\_data\_into\_panel).

Famstand\_grob = "Marital status"

1 = "Married"

2 = "Single"

3 = "Divorced/separated/widowed"

. = "Does not apply/missing information"

educ = "Level of education"

1 = "No degree/general elementary"

2 = "Middle vocational"

3 = "Vocational plus a-levels"

4 = "Higher vocational"

5 = "Higher education"

. = "Does not apply/missing information"

bula = "Federal state of residence"

1 = "Baden-Wuerttemberg"

2 = "Bavaria"

3 = "Berlin"

4 = "Brandenburg"

5 = "Bremen"

6 = "Hamburg"

7 = "Hesse"

8 = "Mecklernburg-Hither Pomerania"

9 = "Lower Saxony"

10 = "North Rhine Westphalia"

11 = "Rhineland-Palatine"

12 = "Saarland"

13 = "Saxony"

14 = "Saxony-Anhalt"

15 = "Schleswig-Holstein"

16 = "Thuringia"

. = "Does not apply/missing information"

emplstat = "Employment status"

1 = "Full-time employment"

2 = "Regular part-time employment"

3 = "Vocational training"

4 = "Marginal, irregular part-time employment"

5 = "Not employed"

. = "Does not apply/missing information"

lfstat = "Labour force status"

1 = "non-working without further information"

2 = "non-working, and older than 65"

3 = "non-working, and currently in a training program"

4 = "non-working, on maternity leave"

5 = "non-working, in military/community service"

6 = "non-working, and registered unemployed"

8 = "non-working, but occasional second job"

9 = "non-working, but working past 7 days"

10 = "non-working, but regular second job"

11 = "working"

12 = "working, but non-working past 7 days"

. = "Does not apply/missing information"

beschaeftigt = "Currently actually employed" (based on the variable lfstat)

1 = "working" (lfstat=11)

0 = "not working" (lfstat=1/10 12)

. = "Does not apply/missing information"

jobposition = "Occupational status"

1 = "Not employed"

2 = "Pensioner"

3 = "Currently in education"  
4 = "Military / community service"  
5 = "Registered unemployed"  
6 = "Apprentice"  
7 = "Self-employed"  
8 = "Manual laborer"  
9 = "Employee"  
10 = "Civil service"  
. = "Does not apply/missing information"

rentner = "Retired"

1 = "Retired"  
0 = "Not retired"  
. = "Does not apply/missing information"

selbstst = "Self-employed"

1 = "Self-employed"  
0 = "Not self-employed"  
. = "Does not apply/missing information"

firmtenure = "Firm tenure (years)"

. = "Does not apply/missing information"

beruf = "Occupation"

0 = "Military"  
1 = "Public administration"  
2 = "Researchers"  
3 = "Technicians, other non-technical occupations"  
4 = "Office employees, commercial employees"  
5 = "Service employees, sellers"  
6 = "Agricultural and fishery professionals"  
7 = "Craftsmen and other similar occupations"

8 = "Operators, assemblers"

9 = "Labourers"

. = "Does not apply/missing information"

sector\_ost = "Sector classification"

1 = "Agriculture, forestry, fishery"

2 = "Mining and quarrying"

3 = "Manufacturing"

4 = "Electricity, gas and water supply"

5 = "Construction"

6 = "Wholesale, retail trade; repair of motor vehicles, motorcycles"

7 = "Hotels and restaurants"

8 = "Transport, storage, communication"

9 = "Financial intermediation"

10 = "Other business services"

11 = "Public administration, defense, social insurance"

12 = "Other public services, private households"

. = "Does not apply/missing information"

firmsize = "No. of employees of the firm person is working for"

1 = "Less than 20"

2 = "20 to 199"

3 = "200 to 1999"

4 = "2000 or more"

. = "Does not apply/missing information"

firmloceast = "Location of the firm in Eastern Germany"

1 = "Located in Eastern Germany"

0 = "Located in Western Germany"

. = "Does not apply/missing information"

jobexp = "Job experience in full-time employment (years)"

. = "Does not apply/missing information"

regalo = "Officially registered unemployed"

1 = "Officially registered unemployed"

0 = "Not officially registered unemployed"

. = "Does not apply/missing information"

nochildren = "Number of children living in the household"

. = "Does not apply/missing information"

nochildren14\_ = "Number of children aged under 14 living in the household"

. = "Does not apply/missing information"

labgro = "Gross labour earnings in the previous month (in DM)"

. = "Does not apply/missing information"

workhours\_week = "Weekly working hours"

. = "Does not apply/missing information"

cpi\_ost = "Consumer price index for Eastern Germany (base year 1995)"

. = "Does not apply/missing information"

labgro\_defl = "Labour earnings deflated using cpi\_ost (base year 1995)"

. = "Does not apply/missing information"

lzf = "Life satisfaction"

scale from 0 = "very unsatisfied" to 10 = "very satisfied"

. = "Does not apply/missing information"

relint = "Realised interview"

1 = "Actually realised interview in the respective wave"

. = "no realised interview in the respective wave"

pidyes = "Party identification present or not"

1 = "Person indicates party identification"

0 = "Person does not indicate party identification"

. = "Does not apply/missing information"

rightpid = "Identification with a right-wing political party"

1 = "Person indicates identification with a right-wing political party"

0 = "Person does not indicate identification with a right-wing political party"

. = "Does not apply/missing information"

leftpid = "Identification with a left-wing political party"

1 = "Person indicates identification with a left-wing political party"

0 = "Person does not indicate identification with a left-wing political party"

. = "Does not apply/missing information"

#### Variables for time-invariant characteristics of individuals in 1990

labgro90 = "Monthly labour earnings in 1990"

selbstst90 = "Individual self-employed in 1990"

firmtenure90 = "Firm tenure in 1990"

jobexp90 = "Job experience in 1990"

educ90 = "Level of education in 1990"

beruf90 = "Job position in 1990"

famstand\_grob90 = "Marital status in 1990"

nochildren90 = "Number of children in 1990"

nochildren14\_90 = "Number of children aged less than 14 in 1990"

sector\_ost90 = "Sector in 1990"

emplstat90 = "Employment status in 1990"

age90 = "Age in 1990"

above40 = "Aged above 40 in 1990"



firmsize90 = "Firm size in 1990"

treat\_dummy\_highlow = "Treatment dummy comparing individuals employed in booming vs. declining industries in 1990"

0 for individuals employed in Agriculture, forestry, fishery; Mining and quarrying; Manufacturing; Electricity, gas and water supply in 1990;

1 for individuals employed in Construction, Hotels and restaurants; Finance, insurance; Other business services in 1990

sector\_change\_rel8993 = "Treatment dummy with relative employment change comparing 1989 and 1993"

#### Further dependent variables

workhours\_month = "Monthly working hours"

log\_stundenlohn\_defl = "Log hourly wage"

not\_employed = "Person currently not employed"

1 = "Not employed"

0 = "Employed"

sector\_stayer = "Sector stayer or mover"

1 = "In the same sector in 1990 and 1993"

0 = "In different sectors in 1990 and 1993"