

## European Data Watch

This section will offer descriptions as well as discussions of data sources that may be of interest to social scientists engaged in empirical research or teaching courses that include empirical investigations performed by students. The purpose is to describe the information in the data source, to give examples of questions tackled with the data and to tell how to access the data for research and teaching. We will start with data from German speaking countries that allow international comparative research. While most of the data will be at the micro level (individuals, households, or firms), more aggregate data and meta data (for regions, industries, or nations) will be included, too. Suggestions for data sources to be described in future columns (or comments on past columns) should be sent to: Joachim Wagner, University of Lueneburg, Institute of Economics, Campus 4.210, 21332 Lueneburg, Germany, or e-mailed to (wagner@uni-lueneburg.de).

### **The Lower Saxonian Salary and Wage Structure Survey – Linked Employer-Employee Data from Official Statistics**

By Gesine Stephan, Universität Hannover\*

During the last decade a new strand of literature has evolved, using linked employer-employee-data sets to attain a better understanding of the interaction between employers and employees in the labor market (Hamermesh 1999). Abowd / Kramarz (1999) summarize the existing research in the new volume of the Handbook of Labor Economics. One focus of their overview is on the data sets underlying the previous studies. For Germany they refer to two data sets only, both of administrative origin. The first is a representative cross-section of firms with representative data on workers, the official Lower Saxonian Salary and Wage Structure Survey (“Niedersächsische Gehalts- und Lohnstrukturerhebung”), surveyed by the Lower Saxonian Statistical Office (“Niedersächsisches Landesamt für Statistik”). The

---

\* I am grateful to Joachim Wagner for helpful hints and to H.-J. Christoff, O. von Holtz, U. Rode and W. Weidlich (Lower Saxonian Statistical Office) for their help in working with the Lower Saxonian Wage Structure Survey.

second is a representative matched worker-firm panel, the IAB Employment Subsample (“Beschäftigtenstichprobe”), containing official data from the Federal Employment Services (“Bundesanstalt für Arbeit”) and administered by the Institute for Employment Research (“Institut für Arbeitsmarkt- und Berufsforschung”).

The data set represented below is the former one, the Lower Saxonian Salary and Wage Structure Survey (an introduction to the IAB Employment Subsample can be found in Bender/Haas/Klose 2000). At first the sampling design and the information included in the Survey are characterized. Subsequently the access to the data is described and selected topics studied with the data set are sketched. Finally further research possibilities using the data are outlined.

### Sampling design

The official German Salary and Wage Structure survey is embedded in the European statistical program. It is conducted in varying periods by the Statistical Offices of the German Federal States, and matched and analyzed for Germany as a whole by the Federal Statistical Office. While there was no survey during the eighties, the subsequent ones refer to 1990 and 1995 and the next year under review will be 2001. Information about organization and methods can be found in Dresch/Kaukewitsch (1993) and Kaukewitsch (1998).

The data are drawn as a two-stage random sample from all establishments with at least 10 employees in the entire manufacturing sector and parts of the service sector (wholesale, retail, finance, insurance). At the first stage establishments, at the second stage employees from these establishment are sampled. The sampling of firms is stratified by industry, region and firm size. Thus, only a few establishments are surveyed in consecutive surveys. The percentage of employees selected per firm depends on firm size and varies between 100 percent for firms with less than 20 employees and 6.25 percent for firms with at least 1000 employees. Even if the same firms are chosen in consecutive surveys it can be expected that the employees included in the data set differ. Only workers covered by statutory pension insurance, not in apprenticeship training and employed during the whole reference year are considered. White-collar workers with a monthly wage rate above DM 17.000 in 1990 and DM 25.000 in 1995 are excluded from the survey.

This procedure covers around 6 percent of the West German and 9 percent of the East German workforce. The 1990 and 1995 data for Lower Saxony include about 65.000 employees from 1.500 firms each year (Statistische Berichte Niedersachsen N I/S- j/90 und N I/S- j/95).

### Information in the survey

Critical for a linked employer-employee data set is the information, which employees work in which enterprise. Additional information on firms is rather rare. Most variables are collected on the individual level, with the focus on compensation. The variables available are summarized in Table 1; they are mostly self-explaining. The information on education and occupation are transmitted by the employer from the social insurance card. The status variable is generated by the Statistical Office based on the pay group and describes roughly the job tasks as they are codified in the collective agreement.

*Table 1*

#### Variables in the Salary and Wage Structure Surveys

##### *Firm level information*

Firm size (number of male and female blue- and white-collar workers)

Sectoral affiliation

Region

Weighting factors

##### *Individual level information*

###### Compensation

October compensation

Gross compensation

Overtime compensation

Taxes and social security contributions

Annual compensation

Gross compensation including bonuses

Net compensation

Annual bonus payments

Method of pay (time pay, piece rate, premium wage)

###### Working time

Contractual weekly working hours

October paid hours

October overtime hours

Working time schedule (working in shifts, on Sundays or at night)

###### Individual characteristics

Gender

Age

Tenure

Education (highest degree up to university level)

Occupation

Status

### **Data access**

To get access to the Lower Saxonian Salary and Wage Structure Survey an agreement with the Lower Saxonian Statistical Office is needed. The recent work with the data is part of a cooperation program between the University of Hannover, the Lower Saxonian Statistical Office and the Lower Saxonian Department of Economics, Technology and Transportation. The critical point is to secure strict confidentiality of the data.

The access procedure is the same as described by Wagner (2000) for the official Lower Saxonian Firm Panel Data: The data are stored on an isolated computer owned by the University of Hannover and located in the office of an Statistical Office employee. Statistical programs written in the University are sent to the Office. The Office runs the programs, checks the output and sends it back to the university if no problems of confidentiality occur.

Regrettably, due to security concerns no test data set containing data from the original Survey is available. I created a random test data set, but since results based on this data set are absolutely meaningless, it does not make much sense to use it in applied econometrics courses.

### **Selected studies with the data set**

What can be done with the data? Most interesting is that they offer the opportunity to control for the firm resp. unobserved firm characteristics through conducting within-firm estimates. In the following some recent studies using the Lower Saxonian data set and exploring its nature as a linked-employer-employee data set are briefly characterized (see Stephan 1999 for a more detailed description):

- The cross-section data sets allow the decomposition of wages in components due to human capital, global firm wage differentials and residuals (Stephan 2001). The global firm wage differentials measure to which amount the wages in a certain firm differs on average from the expected wages, given the observable characteristics of the employees. The estimates reveal, that these firm wage differentials have a large impact on wage differences across blue-collar workers.
- The importance of determinants of annual bonus payments is analyzed using the same methods (Jirjahn / Stephan 1999). It is investigated, how much of the variance of bonus payments is explained through observable human capital and through the employer. The results show that the determination of bonus payments differs from the determination of wages and that the firm explains more (while human capital explains less) of the variance of bonus payments than of the variance of wages.

- The data can be utilized to explore the sources of the gender wage gap, especially the contribution of segregation due to occupation and due to firms (Stephan 1997). The main conclusion is that controlling for observable human capital one third of the remaining gender wage gap in Germany arises from segregation of women in certain occupations while segregation of women in certain firms does not contribute much to wage differences between men and women.

### A look ahead

France is the European country most generous in sharing data from Official Statistics with outside researchers. The French Salary and Wage Structure Survey (“Enquête sur la Structure des Salaires”) has been analyzed by Kramarz/Lollivier/Pelé (1996) and Abowd/Kramarz/Margolis/Troske (1998) in two studies, which provide some inspiration for further analysis based on the German data. In the former one, which uses data for 1986 and 1992, firms are identified that are incorporated in both consecutive surveys. This allows the investigation of changes in the wage policy of firms. The second study matches the Wage Structure Surveys with Firm Panel data, among other things adding information on value-added and sales. Consequently, the relationship between the wage policy of firms and its productivity and profitability can be explored.

Thus, two questions of special interest for further research with the German Wage Structure Survey arise:

1. Is it possible to match data from firms surveyed in two subsequent waves of the Wage Structure Survey?
2. Is it possible to match the Wage Structure Survey data with Firm Panel data (Wagner 2000) available in the Statistical Offices of the German Federal States?

The answer is “yes” for the first question regarding the data from 1995 and the next wave scheduled for the year 2001; enterprise identifiers have not been preserved for the survey from 1990. It is “technically yes” for the second question, since there are still legal restrictions on matching different data sets.

Moreover, the Salary and Wage Structure Surveys are a rich data source for comparisons of the wage structure firstly across the German Federal States and secondly across European countries, since the surveys have to be conducted by all Federal States and in all EU-countries. A first analysis exploiting the potential of the data for international comparisons has been

conducted by Rouault/Kaukewitsch/Söll (1998), who estimate and compare traditional wage regressions for France and Germany. Regarding comparisons across the German Federal States the Lower Saxonian data set has been the only one utilized by outside researchers yet. Thus I would be glad to exchange experiences with researchers trying to get access to the Salary and Wage Structure Surveys sampled by other German Federal States.

### References

- Abowd, J. M. / Kramarz, F.* (1999), The Analysis of Labor Markets using Matched Employer-Employee Data, in: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, Vol. 3b, Amsterdam: Elsevier, 2629 – 2710
- Abowd, J. M. / Kramarz, F. / Margolis, D. N. / Troske, K. R.* (1998), The Relative Importance of Employer and Employee Effects on Compensation: A Comparison of France and the United States, mimeo
- Bender, S. / Haas, A. / Klose, C.* (2000), The IAB Employment Subsample 1975 – 1995, *Schmollers Jahrbuch* 120, 649 – 662
- Dresch, A. / Kaukewitsch, P.* (1993), Methode und Organisation der Gehalts- und Lohnstrukturerhebung 1990, *Wirtschaft und Statistik* 12 / 1993, 879 – 887
- Hamermesh, D. S.* (1999), LEEping into the Future of Labor Economics: The Research Potential of Linking Employer and Employee Data, *Labour Economics* 6, 25 – 41
- Jirjahn, U. / Stephan, G.* (1999), Betriebliche Sonderzahlungen – Eine theoretische und empirische Analyse für die Bundesrepublik, in: Frick, B., Neubäumer, R., Sesselmeier, W. (Eds.), *Die Anreizwirkungen betrieblicher Zusatzleistungen*, München / Mering: Rainer Hampp Verlag, 33 – 68
- Kaukewitsch, P.* (1998), Ergebnisse der Gehalts- und Lohnstrukturerhebung 1996 für 1995, *Wirtschaft und Statistik* 1 / 1998, 46 – 59
- Kramarz, F. / Lollivier, S. / Pelé, L.-P.* (1996), Wage Inequalities and Firm-Specific Compensation Policies in France, *Annales d'Économie et de Statistique* 41 / 42, 2369 – 386
- Rouault, D. / Kaukewitsch, P. / Söll, H.* (1998), Verdienststruktur in Frankreich und Deutschland 1995 im Vergleich, *Wirtschaft und Statistik* 11 / 1998, 867 – 881
- Statistische Berichte Niedersachsen N I/S- j/90, Arbeiter- und Angestelltenverdienste in der gewerblichen Wirtschaft. Ergebnisse der Gehalts- und Lohnstrukturerhebung 1990, Niedersächsisches Landesamt für Statistik
- j/95, Arbeiter- und Angestelltenverdienste in der gewerblichen Wirtschaft. Ergebnisse der Gehalts- und Lohnstrukturerhebung 1995, Niedersächsisches Landesamt für Statistik
- Stephan, G.* (1997) Eine empirische Analyse der Lohnunterschiede zwischen Frauen und Männern, *Statistische Monatshefte Niedersachsen* 51, 5 – 11

- (1999), Die niedersächsische Gehalts- und Lohnstrukturerhebung als Basis arbeitsökonomischer Auswertungen, in: Statistisches Bundesamt (Hrsg.), Möglichkeiten einer wissenschaftlichen Nutzung von Unternehmensdaten aus der amtlichen Statistik, Schriftenreihe Spektrum Bundesstatistik, Band 14, 104 – 117
  - (2001), Firmenlohndifferenziale, Frankfurt / New York: Campus
- Wagner, J.* (2000), Firm Panel Data form German Official Statistics, Schmollers Jahrbuch 120, 143 – 150