

Personnel data of a Russian firm operating in the sector of machine and metal works:

Overview

The personnel data covers 17 years (1990-2006) of a large Russian firm operating in the sector of machine and metal works that produces mainly equipment for oil and gas production. The firm was established in the 1950s and privatized in 1992. The firm successfully managed the transformation from military production to the production of civilian goods. The personnel data was collected through the initiative of Professor Hartmut Lehmann and funded by the German Research Community (DFG) and the European Union (EU).

The data provide detailed information on workers' demographic and human capital characteristics, as well as job-related characteristics, which allow for comprehensive analyses of the internal labor market outcomes. The data file provided for researchers was generated by Ben Kriechel and Karolina Goraus-Tańska, based on two types of source data:

1. *Information from the personnel registry* that contains information on individuals' characteristics such as gender, age, marital status, number of children, education, as well as complete work history.
2. *Information on wages and hours worked* that is collected on a monthly basis.

In the first data file, each row corresponds to a unique worker involving information on individual characteristics and work history. Hence, each observation contains information on one worker.

The second dataset consists of multiple files with information on wages, hours worked, hours of absence, overtime hours, bonus payments, and wage arrears. The dataset has a panel structure, in which every worker is observed on a monthly basis. Hence, each observation contains information on one worker for a month.

While the data for analysis have been anonymized, the raw data are not made publicly available because of confidentiality concerns.

1. Information from personnel registry.

The first column of this data file contains a worker's unique identifier, which allows merging it with the data on wages and hours, and subsequent columns contain information on:

- Demographic characteristics: Date of birth, sex, place of living (i.e., urban/rural or exact name of city or village), marital status (i.e., single, married, divorced, widowed), year of marriage, number of children, date of birth of each child (as separate variables), date of birth of the youngest child; education information including the year of graduation, education level (i.e., basic general, basic professional, secondary general, secondary professional, higher, higher incomplete), educational institution (the last school/university attended), education history (e.g., graduation year, education level and educational institution for every stage of education attended, besides the last one)

ID	birthday	sex	residence	education	...
1	01/10/1960	M	city_name1	higher	...
2	05/12/1971	F	city_name2	basic general	...

