## 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	М	city_name1	higher	
2	05/12/1971	F	city_name2	basic general	

- Work history:
  - Entry date (i.e., date when worker got hired in the analysed firm), exit date (i.e., date when worker left the company, if applicable), the reason why working relationship ended (i.e., retirement, quit, individual dismissal, collective dismissal)
  - Information on every position in which worker was employed: Starting date of employment in the position, unit where employed, position name (i.e., department, function, level in hierarchy, special working conditions), final date of employment in the position. These pieces of information are provided separately for each position level from the first to the last one.
  - Information on every training received inside and outside the firm (if available): Starting date of the training, training name, final date of the training, the location of the training (i.e., inside or outside the firm), schedule of the training (i.e., during or outside the working hours). These pieces of information are provided separately for each training received from the first to the last one.
  - Complete information on prior work history (if available): Starting date of the job before entering the firm, name/type of employer, name of the position, the final date of the job before entering the firm. These pieces of information are provided separately for each job from the first to the last one where the employee worked before entering the analyzed firm.

ID	entry_date	exit_date	exit_reason	start_pos1	unit	position
1	10/03/1998	20/06/2005	retirement	10/03/1998	financial	junior
						accountant

end_pos1	start_pos2	unit	position	end_pos2	start_pos3	
15/05/1999	16/05/2001	financial	senior	19/02/2003	20/02/2003	
			accountant			

## 2. Information on wages and hours worked.

There is monthly information on hours of work and wages including the date of salary payment, wage received and its components (e.g., contractual wage, payment for overtime hours, monthly bonus, yearly bonus, the premium for results, any other types of bonuses, wage arrears), hours worked in a given month, hours of absence, the reason for absence (e.g., annual leave, sickness leave), overtime hours, hours worked during holidays, hours worked during nightshifts.

ID	salary_date	wage_basic	wage_overtime	hours_total	overtime	hours_absence	
1	01/02/2000	310	40	170	20	10	
1	01/03/2000	320	0	160	0	0	
1	01/04/2000	320	20	170	10	0	
1	01/05/2000	310	0	150	0	10	
1	01/06/2000	350	0	160	0	0	
2	01/03/1999	100	0	80	0	0	
2	02/03/1999	100	0	80	0	0	