Data for 'Status and Just Gender Pay Gaps. Results of a Vignette Study'

Kinga Wysieńska-Di Carlo and Zbigniew Karpiński

The dataset

The data come from a vignette survey carried out in Poland in 2017 on a nation-wide sample of 2,001 men and women aged 20 to 64. The sample was not probabilistic, but it was designed to match the target population on a number of characteristics, such as gender, age, degree of education, or region.

The data come in three formats: R, Stata, and SPSS. The three corresponding files have the same name and the same structure.

Each dataset contains the following variables:

- `respid` — respondent's identifier;
- `vignid` — vignette's identifier;
- `r_gender` — respondent's gender, with categories "male" and "female";
- `just` — the amount of earnings (in the Polish currency, PLN) that respondent with ID `respid` thinks is just for vignette with ID `vignid`;
- `lnjust` — logged just earnings, or the natural logarithm of `just`;
- `v_psco` — vignette's occupation coded using a 4-digit code from the Polish Sociological Classification of Occupations (PSCO) by Domański, Słomczyński, and Sawiński (2009);
- `v_occ` — PSCO title of vignette's occupation;
- `v_isco` — vignette's occupation re-coded using a 4-digit code from International Standard Classification of Occupations 2008 (ISCO-08);
- `v_isco3` — a 3-digit ISCO-08 category of vignette's occupation;
- `v_prscale` — PSCO prestige score of vignette's occupation;

- `v_prcat` — occupational prestige category of vignette's occupation: (*a*) high, for occupations with a prestige score over 70; (*b*) medium, for occupations with a prestige score higher than 40, but not higher than 70; and (*c*) low, for occupations with a prestige score up to 40;
- `v_high` — a dummy variable coded 1 for high prestige occupations and 0 otherwise;
- `prop_fem` — proportion of women in vignette's occupation (defined at the level of ISCO-08 3-digit category);
- `v_segr` — classification of vignette's occupation into *male-dominated* (with more than 70% male incumbents), *female-dominated* (with less than 30% male incumbents) and *gender-balanced*;
- `v_gender` — vignette's gender, with categories "male" and "female";
- v_age — vignette's age (in years);
- `v_agebin` — a dummy variable coded 1 for vignettes aged over 40 and 0 otherwise;
- `v_child` — the number of vignette's children;
- `v_childcat` — the number of vignette's children divided into 3 categories: (*a*) no children; (*b*) 1 or 2 children; and (*c*) 3 or more children;
- `v_jobexp` — vignette's years of job experience.

References

Domański, Henryk, Kazimierz M. Słomczyński, and Zbigniew Sawiński. 2009. *Sociological Tools Measuring Occupations: New Classification and Scales.* Warsaw: IFiS Publishers.