The dataset contains estimates of gender wage gap (GWG) for hourly wages using various methodological approaches for multiple countries over time.

Data source employed for estimations of GWG: EU-SILC, waves 2004-2013.

For each country and year, sample was limited to persons aged 16-65 that were employed full-time during the last 12 months, had only one job, and the following other variables were available: information on yearly income from dependent employment and hours usually worked, age, gender, marital status, education, place of residence, occupation, industry, firm size, presence of children in the household, presence of other earner in the household, presence of non-earned income in the household.

Variables:

* *country:* string variable indicating for which country the GWG estimate applies. Countries covered: AT (Austria), BE (Belgium), BG (Bulgaria), CZ (Czechia), DE (Germany), DK (Denmark), EE (Estonia), ES (Spain), FI (Finland), FR (France), GR (Greece), HU (Hungary), IT (Italy), LT (Lithuania), LV (Latvia), NL (Netherlands), PL (Poland), PT (Portugal), RO (Romania), SE (Sweden), SI (Slovenia), SK (Slovakia), UK (United Kingdom);
* *year*: numerical variable indicating for which year the GWG estimate applies. Variable takes values between 2003 and 2012, as the EU-SILC waves 2004-2013 for most of mentioned countries contain information for 2003-2012. This is due to the fact that income reference period is the previous calendar year. The only exception among mentioned countries is the United Kingdom where income reference period is the current year. For most countries GWG estimates for years 2004-2012 are available. For Austria, Belgium, Denmark, Estonia, France and Sweden estimates are available already for 2003. For the United Kingdom estimates are available from 2005. For Bulgaria, Latvia, and Romania estimates are available from 2006.
* *method*: string variable indicating the method used for estimation of the GWG, and point of the distribution at which the gap was estimated (if different than the mean). “blin” stands for Blinder (1973) method, “oaxa” for Oaxaca (1973), “sloc” for Słoczyński (2013), “cott” for Cotton (1988), “reim” for “Reimers (1983), “fort” for Fortin (2008), and “neum” for Neumark (1988). Those are parametric methods that differ regarding the counterfactual wage assumed. “hblin”, “hoaxa”, “hsloc”, “hcott”, “hreim” stand for parametric approaches with Heckman (1979) correction for selection from active into full-time full-year employment (information on presence of children/other earner/non-earned income in the household used as exclusion criteria). “jmp25”, “jmp50”, and “jmp75” stand for Juhn, Murphy, and Pierce (1993) method at the 25th, 50th, and 75th percentile of wage distribution. In each case multiple counterfactual wage structures are assumed, thus the full names are e.g. “jmp25\_blin”. The same logic applies to names standing for Recentered Influence Functions, “rif”, method of Firpo, Fortin, and Lemieux (2009) that are also estimated at 25th, 50th, and 75th percentiles and with multiple counterfactual wage structure assumed. “dflm”, “dfl25”, “dfl50”, and “dfl75” stand for estimates with method of DiNardo, Fortin, and Lemieux (1996) at the mean, and at the 25th, 50th, and 75th percentile. “nopo” stands for Nopo (2008). In reweighting method of DiNardo, Fortin, and Lemieux (1996), and matching decomposition of Nopo (2008) the males’ wage structure is treated as non-discriminatory.
* *set:* methods from the variable *methods* were applied to different sets of controlled characteristics. The following sets were distinguished: "basic" stands for set of demographic and human capital variables only, and includes controls for age, education, marital status, and the degree of urbanisation in the place of living; "ind" stands for basic set plus the controls for the industry, "ind2" stands for set of characteristics that covers basic set plus industry and firm size controls; "occ" covers basic characteristics plus occupation dummies; "indocc" stands for set of controls for basic characteristics plus industry and occupation controls; "ind2occ" is the set that covers all available characteristics.
* *gwg:* numerical variable that indicates the estimate of the adjusted gender wage gap for given country, in given year, under given methodological approach, and with given set of characteristics controlled. For each country in given year, there are 318 estimates of the GWG.

**References**

Blinder, A. S., 1973. Wage discrimination: Reduced form and structural estimates. Journal of Human Resources 8 (4), 436–455.

Cotton, J., 1988. On the decomposition of wage differentials. The Review of Economics and Statistics, 236–243.

DiNardo, J., Fortin, N. M., Lemieux, T., 1996. Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. Econometrica 64 (5), 1001–44.

Fairlie, R.W., 1999. The absence of the African-American owned business: An analysis of the dynamics of self-employment. Journal of Labor Economics 17 (1), 80–108.

Fairlie, R. W., 2005. An extension of the Bblinder-Oaxaca decomposition technique to logit and probit models. Journal of Economic and Social Measurement 30 (4), 305–316.

Firpo, S., Fortin, N. M., Lemieux, T., 2009. Unconditional Quantile Regressions. Econometrica 77 (3), 953–973.

Fortin, N., Lemieux, T., Firpo, S., 2011. Decomposition methods in economics. In: Handbook of Labor Economics. Elsevier, pp. 1–102.

Fortin, N. M., 2008. The Gender Wage Gap among Young Adults in the United States: The Importance of Money versus People. Journal of Human Resources 43 (4).

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–61.

Juhn, C., Murphy, K. M., Pierce, B., 1993. Wage Inequality and the Rise in Returns to Skill. Journal of Political Economy 101 (3), 410.

Lemieux, T., June 2006. Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? American Economic Review 96 (3), 461–498.

Machado, J. A., Mata, J., 2005. Counterfactual decomposition of changes in wage distributions using quantile regression. Journal of applied Econometrics 20 (4), 445– 465.

Neumark, D., 1988. Employers’ discriminatory behavior and the estimation of wage discrimination. Journal of Human Resources, 279–295.

Nopo, H., 2008.Matching as a tool to decompose wage gaps. The Review of Economics and Statistics 90 (2), 290–299.

Oaxaca, R., 1973. Male-female wage differentials in urban labor markets. International Economic Review 14 (3), 693–709.

Oaxaca, R. L., Ransom, M. R., 1988. Searching for the effect of unionism on the wages of union and nonunion workers. Journal of Labor Research 9 (2), 139–148.

Oaxaca, R. L., Ransom, M. R., 1994. On discrimination and the decomposition of wage differentials. Journal of Econometrics 61 (1), 5–21.

Reimers, C. W., 1983. Labor market discrimination against Hispanic and Black men. The Review of Economics and Statistics 65 (4), 570–579.

Słoczynski, T., 2015. The Oaxaca–Blinder unexplained component as a treatment effects estimator. Oxford Bulletin of Economics and Statistics 77 (4), 588–604.