

**Basic Oaxaca-Blinder decomposition**

In the lecture, a decomposition of the gender wage gap (log wages) was shown using data from the GSOEP, where schooling and experience (quadratic effect) served as covariates (see slide 18 for data preparation; slides 28/29 for the decomposition). The data can be found on Ilias (`gsoep-extract.dta`). Using the same data setup, do the following.

1. Extend the  $X$  variables of the model by tenure (number of years working for the current employer; variable `tenure`), the occupational status as measured by the international socio-economic index (`isei`), and the number of children under the age of 16 in the household (`children`). Also take account of the survey design, i.e. primary sampling units (`psu`), sampling weights (`weight`), and strata (`strata`).

Hint: See `help svyset` for information on how to set the survey design.

2. Compute the aggregate and detailed Oaxaca-Blinder decomposition. How do the results change compared to the specification used in the lecture? (How does the survey design change the results? How do the additional variables change the results? Make sure to use the same estimation sample when comparing results.)

Hint: In command `oaxaca` specify option `svy` to take account of the survey design defined by `svyset`. This is somewhat different from most official Stata commands, to which `svy` can be applied as a prefix (i.e. `svy: command`).

3. Confirm the results returned by `oaxaca` (for the extended decomposition including survey design) by computing the aggregate Oaxaca-Blinder decomposition “by hand” (that is, estimate the means of the variables and the regression coefficients and then compute the decomposition from these outputs, and not using `oaxaca`). Also compute the contribution of schooling to the “explained” part and the “unexplained” part by hand.