

Part I: The index problem

1. Using the extended decomposition from Exercise 1 (i.e. schooling, full-time experience, ISEI, number of children), evaluate how the results change depending on how you handle the index problem. Compute the following variants:
 - male coefficients
 - female coefficients
 - pooled model
 - threefold decomposition
2. Generate an overview table and try to make sense of the results. What is the correct interpretation of the various results? How can the differences be explained?
3. Optional: Compute a decomposition that is defined in a way such that the unexplained component can be interpreted as an “average treatment effect” (see slides for details).

Part II: The transformation problem

1. Replace ISEI in the extended model of Exercise 1 by the (categorical) EGP variable (**egp**). Before you do that, inspect the variable **egp** carefully and possibly drop categories with a very low number of observations. Only report the aggregate contribution of EGP. Illustrate how the results change if you switch the base level.
2. Normalize the effects of EGP to make its contribution independent of the choice of the base level (unweighted normalization using **oaxaca**).
3. Now simplify the EGP variable by combining classes VIIa and VIIb (codes 7 and 8) into one bigger class. How do the decomposition results change?
4. Optional: Compute the contribution of EGP and the simplified EGP to the unexplained part using a weighted normalization. You need to do this manually (hint: you can use command **contrast** to obtain normalized coefficients after running a regression). Compare the results to the results from the unweighted normalization.
5. Optional: Compute the “industry decomposition” described on the slides by economic sector (variable **industry**).