

Due April 17 in class

Individual data are provided for a random sample of workers employed in 2010 in the Retail Trade industry. These data are available from the web page for this class in two formats: retail\_e696i.dta (STATA), and retail\_696i.xlsx (Excel). Data on the variables below are provided. For standard error calculations you may condition on sample means.

Variable	Definition
<i>hrlywage</i>	hourly wage in \$'s per hour
<i>married</i>	= 1 if married (otherwise = 0)
<i>education</i>	number of years of completed schooling
<i>exp</i>	number of years of potential work experience
<i>occ1</i>	=1 if a managerial/professional/technical worker (= 0 otherwise)
<i>occ2</i>	=1 if a clerical worker (= 0 otherwise)
<i>occ3</i>	=1 if a sales worker (= 0 otherwise)
<i>occ4</i>	= 1 if a blue collar worker (= 0 otherwise)

Consider the following wage models for workers in retail trade:

$$\ln(\text{hrlywage}_i) = \beta_0 + \beta_1 \text{married}_i + \beta_2 \text{education}_i + \beta_3 \text{exp}_i + \beta_4 (\text{exp}_i)^2 + u_i \quad (1)$$

$$\ln(\text{hrlywage}_i) = \beta_0 + \beta_1 \text{married}_i + \beta_2 \text{education}_i + \beta_3 \text{exp}_i + \beta_4 (\text{exp}_i)^2 + \sum_{j=1}^3 \beta_{4+j} \text{occ}j + u_i \quad (2)$$

1. Estimate equations (1) and (2) separately for males and females.

- (a) Conduct gender wage decompositions in logs assuming that the wage structure for males is the non-discriminatory standard, i.e.

$$\ln(G + 1) = \ln(D + 1) + \ln(Q + 1)$$

where  $G$ ,  $D$ , and  $Q$  are respectively the gross/raw gender wage gap, discrimination (unexplained gap), and qualification (explained gap)

- (b) Estimate the values of  $G$ ,  $D$ , and  $Q$  along with the standard errors on your estimates of  $D$  and  $Q$  using the delta method.

(Question 2 on the next page)

2. Estimate equations (1) and (2) for the combined (pooled) samples of males and females.

(a) Conduct generalized decompositions of the gender wage gap in logs, i.e.

$$\ln(G + 1) = \ln(\delta_{mo} + 1) + \ln(\delta_{of} + 1) + \ln(Q + 1)$$

where  $\delta_{mo}$  and  $\delta_{of}$  are respectively favoritism towards males and pure discrimination.

(b) Estimate the values of  $\delta_{mo}$ ,  $\delta_{of}$ , and  $Q$  along with the standard errors on your estimates of  $\delta_{mo}$ ,  $\delta_{of}$ , and  $Q$  using the delta method.