Problem Set II

EE325 Section 1

1) Use the data in 401KSUBS.DTA for this question, restricting the sample to *fsize* = 1
   a. To the model estimated in Table 8.1, add the interaction term, *e401k*inc.* Estimate the equation by OLS, then obtain the usual and robust standard errors. What do you conclude about the statistical significance of the interaction term?
   b. Now estimates the more general model by WLS using the same weights, $1/inc_i$, as in Table 8.1. Compute the usual and robust standard error for the WLS estimator. Is the interaction term statistically significant using the robust standard error?
   c. Discuss the WLS coefficient on *e401k* in the more general model. Is it of much interest by itself? Explain
   d. Re-estimate the model using WLS, but use the interaction term *e401k*(inc-30);* the average income in the sample is about 29.44. Now interpret the coefficient on *e401k*.

TABLE 8.1 Dependent Variable: *nettfa*

Independent Variables	(1) OLS	(2) WLS	(3) OLS	(4) WLS
inc	.821 (.104)	.787 (.063)	.771 (.100)	.740 (.064)
$(age - 25)^2$	—	—	.0251 (.0043)	.0175 (.0019)
male	—	—	2.48 (2.06)	1.84 (1.56)
e401k	—	—	6.89 (2.29)	5.19 (1.70)
intercept	−10.57 (2.53)	−9.58 (1.65)	−20.98 (3.50)	−16.70 (1.96)
Observations	2,017	2,017	2,017	2,017
R-squared	.0827	.0709	.1279	.1115

2) Use the data in MINWAGE.DTA for this exercise, focusing on sector 232.
   a. Estimate the equation;

$$gwage232_t = \beta_0 + \beta_1 gmwage_t + \beta_2 gcpi_i + u_t$$

   And test the errors for AR(1) serial correlation. What do you conclude overall?
   b. Obtain the Newey-West standard error for the OLS estimation in part (i), using a lag of 12. How do the Newey-West standard errors compare to the usual OLS standard errors?
   c. Now obtain the heteroskedasticity-robust standard errors for OLS, then compare them with the usual standard errors and the Newey-West standard errors. Does it appear that serial correlation or heteroskedasticity is more of a problem in this application?
   d. Use the Breusch-Pagan test in the original equation to verify that the errors exhibit strong heteroskedasticity.