OLS Estimation

The OLS estimation of the model with GPA as dependent variable and gender and preparatory course participation as explanatory variable are the same as in the lecture with a coefficient of -0.21 for gender and a coefficient of around 0.82 for the dummy variable of participating in the preparatory course. The standard errors are around 0.04 and 0.05 respectively and the t-values confirm that both variables are highly significant at a 1 percent level.

Model 1: OLS, using observations 1-1000
Dependent variable: GPA

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	5.77111	0.0339729	169.8739	<0.0001	***
GENDER	−0.213759	0.0443118	−4.8240	<0.0001	***
PARTICIPATION	0.824368	0.0468589	17.5926	<0.0001	***

Mean dependent var	5.944610	S.D. dependent var	0.802589
Sum squared resid	486.4320	S.E. of regression	0.698495
R-squared	0.244090	Adjusted R-squared	0.242574
F(2, 997)	160.9703	P-value(F)	2.61e-61
Log-likelihood	−1058.609	Akaike criterion	2123.219
Schwarz criterion	2137.942	Hannan-Quinn	2128.815

TSLS Estimation

The first stage regression of the TSLS estimation is a regression of the endogenous variable participation on all instruments, i.e. gender and the dummy variable for receiving an email invitation. The fitted values for the variable participation of the first stage are retained for the second regression of GPA on gender and the fitted values for participation obtained from the first stage OLS. The second stage estimation confirms the parameter estimates from the lecture, namely -0.17 for gender and 0.24 for the participation variable.

Model 2: TSLS Stage 1, using observations 1-1000
Dependent variable: **PARTICIPATION**

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.101123	0.0229047	4.4150	<0.0001	***
GENDER	0.0484557	0.0268981	1.8015	0.0719	*
EMAIL	0.412899	0.0269009	15.3489	<0.0001	***

Mean dependent var	0.337000	S.D. dependent var	0.472921
Sum squared resid	179.7294	S.E. of regression	0.424582
R-squared	0.195593	Adjusted R-squared	0.193980
F(2, 997)	121.2115	P-value(F)	7.58e-48
Log-likelihood	−560.7870	Akaike criterion	1127.574
Schwarz criterion	1142.297	Hannan-Quinn	1133.170

Model 3: TSLS Stage 2, using observations 1-1000
Dependent variable: **GPA**

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	5.94787	0.0511694	116.2388	<0.0001	***
GENDER	−0.172762	0.0512149	−3.3733	0.0008	***
PARTICIPATION	0.240499	0.122459	1.9639	0.0498	**
fitted					

Mean dependent var	5.944610	S.D. dependent var	0.802589
Sum squared resid	634.9784	S.E. of regression	0.798053
R-squared	0.013251	Adjusted R-squared	0.011272
F(2, 997)	6.694381	P-value(F)	0.001294
Log-likelihood	−1191.856	Akaike criterion	2389.713
Schwarz criterion	2404.436	Hannan-Quinn	2395.309

Standard Errors from TSLS Stage 2

The standard errors for the final regression are 0.05 and 0.12 for the variables gender and participation respectively. However, these are not the true coefficients because the second stage regression does not account for the first stage regression, i.e. that the participation variable is an estimated regressor. Hence the variance will be wrong and this will also cause that the obtained standard errors from the second stage will be wrong.