(1) Matlab code for this question has been posted on the website. I won’t review solutions for the coding questions; those can be found by inspection of the uploaded m-file.

(1c) The class size coefficients suggest that a one-student increase in class size will increase reading test scores by about .14 points, and math scores by about .22 points. This seems counter-intuitive, as we tend to think that smaller class sizes, where teachers can dedicate more time per student, would lead to better student outcomes.

In terms of mean-independence, it is difficult to justify this. It could be the case that larger class sizes are also correlated with other things - just as an example, schools with larger average class sizes may tend to be larger, and perhaps larger schools can offer more programs to students and are the ones more sought-out by families.

(1e) In both cases, when controls for the percentage disadvantaged and enrollment are added, the coefficient on classsize is reduced, and even becomes negative for the case of reading test scores. The negative coefficient seems more consistent with our prior expectations. Once we have controlled for variation across schools in their size and a measure of family income, the class size effects became smaller.

Although it is now, perhaps, more credible to maintain mean independence than it was for the regression in (1c), it is still a bit tenuous. For example, it could be the case that families who are unobservably dedicated to their child’s education and seek smaller classes may find ways to enroll their children in such schools. Thus, even after controlling for enrollment and percent disadvantaged, the class size variable may still confound a pure class size effect with a family background effect. One could tell lots of similar stories here that question the validity of mean-independence.

(1f-1h) Answers here are found primarily in the code. Note, for the case of (1h):

\[
E(\text{ReadingScore} \mid \cdot, \text{“Before”}) = \beta_0 + \beta_1 \text{ClassSize}_0 + \beta_2 \text{PrctDisadv}_0 + \beta_3 \text{Enroll}_0
\]

and, after, the firm is in place:

\[
E(\text{ReadingScore} \mid \cdot, \text{“After”}) = \beta_0 + \beta_1 [\text{ClassSize}_0 - 10] + \beta_2 [\text{PrctDisadv}_0 - 20] + \beta_3 \text{Enroll}_0.
\]

The object of interest, which we will call \( \Delta \), is the difference between these two:

\[
\Delta \equiv E(\text{ReadingScore} \mid \cdot, \text{“After”}) - E(\text{ReadingScore} \mid \cdot, \text{“Before”}) = -10\beta_1 - 20\beta_2.
\]

This leads to

\[
\hat{\Delta} = -10\hat{\beta}_1 - 20\hat{\beta}_2 \approx 7.22.
\]
To calculate a standard error, we can first define \( H = [0 \quad -10 \quad -20 \quad 0] \) so that
\[
\hat{\Delta} = H \hat{\beta}
\]
which implies
\[
\text{Var}(\hat{\Delta} | X) = H\text{Var}(\hat{\beta})H' = \sigma^2 H(X'X)^{-1}H',
\]
from which a standard error can be calculated.

(2) The slope coefficient should equal one.

To see this note,
\[
\hat{u} = y - X\hat{\beta} - [I - X(X'X)^{-1}X']y = My.
\]
Thus our coefficient estimate is simply
\[
(\hat{u}'\hat{u})^{-1}\hat{u}'y = [y'M'y]^{-1}y'M'y = [y'M'y]^{-1}y'M'y = 1.
\]