M1.

The authors are interpreting the coefficients when the predictor variables are collinear. A very clear example shows up in the juxtaposition of the coefficient values for the predictors describing the proportions of elderly males and females one is positive and one is negative in almost the same magnitude. This makes sense when you think about it. I think it is safe to say that in any population, the proportion to males is negatively correlated to the proportion of females. A positive correlation to a higher proportion of males would mean a negative correlation to the proportion of females since they are opposite.

The appropriate thing to do would have been to perform a Lasso regression to turn off select predictors that are highly correlated with other predictors or a Ridge regression if you are very passionate about keeping the variables in since the proportional age brackets for either gender are not the same.

M2.

When developing a model for theory, researchers are normally concerned with discovering the predictors that directly contribute to the practical problem they are studying. It is the correlation vs causation problem. In this case, variable selection is important because there may likely me predictors that have no causation, but are correlated regardless because they are correlated with a predictor that has causation.

If we were only concerned with prediction and not theory, we would not have as much motivation to find the true cause of problem and can stay satisfied with more crowded predictors.

M3.

Ridge regression solves the issue of multicollinearity by adding bias to the coefficients. There cannot be multicollinearity if there is only one predictor so you have no reason to use it.