This paragraph comes from my DAEN 500 final exam question 5. I chose not to include this answer because I felt it was beyond the scope of the question and also did not know how to plot these in order to “demonstrate” this answer.

Finally, other circumstances which may dictate whether or not to use a multiple linear regression model is based on validating certain assumptions which are required for this model to provide accurate results. One should not use it when these specific assumptions are not valid within the population. These include: (from zyBooks participation activity 23.3.1)

* mean of zero assumption – this assumption relates to mean of the residuals – the distance of the observations from the predictor line. The distance should be approximately zero for each fixed value on the horizontal axis. This is validated using a residual scatterplot with the residuals on the vertical axis and a predictor on the horizontal axis. If the observations are close to the line, then this assumption holds true.
* constant variance assumption – the variance of the residuals should be approximately constant for each fixed value on the horizontal axis. Again, this is validated using a residual scatterplot with the residuals on the vertical axis and a predictor on the horizontal axis.
* independence assumption – the value of one error should be independent from the value of any other error. This assumption is difficult to assess unless time is a variable and the data is available and was collected in time order. However, if able, one can validate this through a scatterplot with residuals on the vertical axis time order on the horizontal axis. If the residuals seem random, then one can conclude that this assumption holds true.
* normality assumption – the errors are assumed to be normally distributed. This can be assessed using a normal probability plot of residuals with the ordered residuals on the vertical axis and theoretical normal quantiles on the horizontal axis. If the plotted points lie reasonably close to the line, then one can conclude that this assumption is valid.

When these assumptions can be validated, then a multiple linear regression model may be used. However, that is not to say that when any of these assumptions are invalid that a logistic regression is necessarily appropriate; only that a multiple linear regression is not.