In my spec, I was assuming that there is nonlinearity existed between the predictor variables and the response variable, but after the analysis, I found that a linear model would give the least residuals, so the best prediction. So, I used linear least squared.

It also results that my original thought about the weighing between maximizing conversion (response variable) and minimizing spent (one of the predictor variables) would just be on a straight line, and the position only depends on the tuning value. Also, the Pareto curve between them would also be meaningless. So instead, I did a classification problem like question 3 in homework 9. I consider ads that have at least one actual conversion (the person became an actual active user) to be one class, and no conversion to be another class.