Elementary Education

As described by the US news and world report, "the elementary years – from kindergarten through third grade – are particularly important ones in children's schooling. Parents and teachers know that children acquire new skills and knowledge rapidly during these years. Research shows that average annual learning gains for children in grades K-2 are dramatically greater than those for subsequent years of school. Moreover, the outcomes of early elementary education, particularly whether or not a child can read proficiently by third grade, are a powerful predictor of later school and life outcomes."

Due to the importance of elementary education, in this analysis you will be looking at how various factors affect learning in elementary school as measured by results on a standardized test. The SchoolResults.txt data contain information on various school districts throughout California and includes the following variables:

- Score: Average cumulative Score on the Stanford 9 standardized test (out of 1600)
- Lunch: Percent qualifying for reduced-price lunch
- Computer: Number of Computers
- Expenditure: Expenditure per student
- Income: District average income (in USD 1,000)
- English: Percent of English learners
- STratio: Student-to-teacher ratio

So far, we have discussed three different modeling paradigms/methods: multiple linear regression, penalized regression, and nonlinear regression. As our methods have gained flexibility/complication, we lose some ability to really explain what's going on in the model. Your first goal is to find a good model (propose two reasonable models and choose the best one). Second, answer the following questions to the extent that your chosen model can. If you cannot answer a question based on your model choice, be sure to clarify how the improved fit/predictive power of the model merits the lack of interpretabilty, or vice versa. Here are the questions of interest:

1. "Income" is generally a measure of how much money a school has to spend on extracurricular activities (as opposed to expenditures which i s how much spent per student in the class room). Is there evidence of diminishing returns on extracurricular activities in terms of student learning?
2. Is English as a second language a barrier to student learning?
3. In your opinion and based on the data, what can be done to increase student learning?
4. (How well does the model predict compared to alternatives?)