CHAPTER 5

CONCLUSIONS

The purpose of this dissertation has been to review statistical methods used for assessing school effectiveness, to assess school effectiveness in Greece using multilevel models and to make adjustments for the previous achievements of the students. Many authors, whose work has been reviewed in this thesis, argue that the use of raw, unadjusted results lead to invalid comparisons and misleading outcomes, as far as the effectiveness of the schools is concerned. This conclusion has been confirmed by our results, too. Furthermore, some interesting differences, with respect to gender, to the type of institution and to the scientific orientation that students have chosen, have been observed.

To be more precise, first of all, if we do not make adjustment for the 3^{rd}-grade score we conclude that girls do much better than boys in the National Entrance Exam in Greece and that the students of the 1^{st}, 2^{nd} and 3^{rd} scientific orientation do better than those in the 4^{th} one. The difference between the 2^{nd} and, especially, the 3^{rd} scientific orientation with the 4^{th} orientation is very large. On the other hand, if we make adjustment for the background of the students, that is the 3^{rd}-grade score, then the results would be very different. It has been observed that the 3^{rd} grade score is a very significant explanatory variable, since the estimate of the standard error of the parameter is less than a third of the parameter estimate (0.784(0.005)). In this case, we concluded that boys do better than girls. This means that boys make more progress than girls in the National Entrance Exam with respect to their 3^{rd}-grade score. It was also concluded that public schools do much better than private ones, while the scientific orientation differences are not so pronounced and that students who choose the 4^{th} scientific orientation do better than ones who choose the first

one. Finally, it was found that the students who took the Exam in 1990 did better than those who took the Exam in 1991.

In order to identify schools or prefectures with residuals which are substantially different we ordered the residuals from smallest to largest and confidence intervals about each residual were constructed. Through this procedure we concluded that the prefecture with the highest mean score, for both years, was the prefecture of Corinthia for the year 1990, the second best prefecture was Attica for the year 1990 while the prefecture of Evros for the year 1991 had the lowest mean score for both years.

Nevertheless, we have to keep in mind that there are limitations in making comparisons between institutions and that when we apply a statistical model we have to treat the results as suggestive rather than definitive (Goldstein, Spiegelhalter (1996)). When comparative information about institutions are to be analyzed, it must be handled sensitively and with regard to all its problems and limitations.