PART III

GENERALIZATIONS OF THE LINEAR MODEL

Having built up the regression edifice affectionately known as the classical model, we turn to critical reconsideration of the elements of the theory. Every assumption represents a restriction on the data-generating process and every data set presents opportunities for unclassical behavior. Exceptions to our restrictions will generally result in exceptions to the statistical properties that we have derived and in possibilities for misguided inferences.

There is a clear hierarchy among the assumptions. The specification of the regression function is foundational. Without it, we would never have proceeded to take up the others. The most recent assumption, the specification of the distribution function as a member of the family of normal distributions, is the most narrow and most dispensible. In the remainder of this book, we will reconsider each of the assumptions in the reverse of the order in which we introduced them, working our way back to the most fundamental components of the theory.

Our analysis will follow a pattern. We begin the review of each assumption with the typical reasons for questioning it. These doubts provide alternative specifications of our model, which often generalize the classical model in some way. Given such generalizations, we then reconsider the properties of the classical methods, checking whether they continue to work and, if not, how they may fail. Failures naturally lead to a search for diagnostic tests to detect each deviation from classical conditions and for alternative estimation methods that do not share the weaknesses of OLS.

As you probably expect, the analysis continues to grow in its complexity. The fundamental departure from the previous material is that we must work with *nonlinear* estimators. Linearity, or sometimes quadraticity, in the dependent variable $\mathbf{y}$ has been a critical characteristic of the statistics that we have studied up to this point. We will frequently encounter statistics that do not possess such convenient analytical forms hereafter. The variety and complexity of these forms are some of the most intimidating characteristics of this new material.

We will cope with this variety and complexity in two basic ways:

1. interpreting many new methods as an approximate application of the OLS method that we have already studied in detail; and
2. showing how the approximate distribution theory is essentially analogous to the exact distribution theory of the classical model.

In fact, such exact results as those that we have been able to provide so far generally elude the analyst in the problems to come. Exact moments and distributions for our statistics are simply not available, except perhaps as numerical calculations for particular experiments. Thus, analysts have sought approximate results and, fortunately, many situations provide a delightfully familiar theory. This theory reproduces, in effect, results that have clear counterparts in the theory of the classical regression model.