This was the case for firm objectives in the Attributes model, and target pest and level of infestation in the Specific Practices models. Statistics on the number and proportion of refusals for each survey question or variable in the analysis are summarized in Table 7.

There were a number of other data/numerical problems which made it necessary to drop variables from the three models. Linear dependencies in the General Practices model were the most troubling. This was probably due to the relatively large number of binary or discrete choice explanatory variables in this model. Deleting at least one of the suspect variables was the only means available to remedy this kind of difficulty. Some dependencies involved more than six variables and could never be completely identified.

Multicollinearity was also a problem in most regressions. Condition numbers were calculated for each regression and ranged from 43.9 in the Specific Practices model for Chlorothalonil, to 730.29 in the General Practices model for strawberries. Serious multicollinearity occurred between a number of binary response variables (with low variance) and the intercept term. For example, more than 94 percent of all growers indicated that they generally used economic thresholds for making insect and fungus pest control decisions. Since more detailed information on this issue was available in the Specific Practices section of the interviews, it was decided to drop these variables from the General Practices models. In the Attributes models, variables which were highly collinear included: (1) acres owned and rented, and gross revenue; (2) firm form and affiliation; (3) different levels of pesticide certification; (4) various components of the weather variables, and; (5) location variables with a wide array of other variables such as firm size and soil type. In the General Practices models, soil-testing frequency and following test recommendations were highly collinear, and so were comparable variables for plant tissue analysis.

One means of mitigating the consequences of multicollinearity is through the application of principal components techniques. Principal components are a set of mutually orthogonal vectors composed from linear combinations of the original regressors. They are derived sequentially in order of decreasing variability. Consequently, the last few principal components will represent only a small fraction of the variability of the original data, or conversely, a significant proportion of the multicollinearity in the data. Thus it is often possible to drop these last few principal components from the model and substantially improve the power of the estimated coefficients. Detailed expositions on principal components techniques can be found in Chatterjee and Price, and Maddala.

The deletion of a subset of principal components as described above, in effect imposes a restriction on the model consistent with the multicollinearity of the data. Of course, in general, any restriction inappropriately imposed on an empirical model will statistically bias the coefficient estimates. To minimize this undesirable effect, conservative guidelines were used to implement these procedures. A sufficient number of components were retained so that at least 95 percent of the variation of the explanatory data was preserved, or F-tests for the restrictions imposed by the dropped components had a type-one error probability in excess of 80 percent. Because of these conservative criteria, principal components procedures were not applied to every regression. Consequently, there were several cases where multicollinearity could not be mitigated.