that the (conditional) distribution of u_i is normal, and following Tobin's (1958) original procedure of estimating the parameters by maximum likelihood. The log likelihood function for this problem is globally concave, so that it is a routine problem in non-linear estimation, typically no more time consuming than the estimation of a probit. Note also that if (25) is correct, OLS will be inconsistent for the parameters β. The regression function is

$$E(y_i|x_i) = (1 - F(-x_i'\beta/\sigma))x_i'\beta + \sigma \int \tau \, dF(\tau) \qquad (26)$$

where $F(.)$ is the distribution function of $\sigma^{-1}u_i$. (26) will generally not be linear in x_i.

Tobin's maximum likelihood method works well when its assumptions are satisfied. However, the estimates will typically be inconsistent if normality fails, or perhaps more seriously, if there is heteroskedasticity [see Arabmazar and Schmidt (1981, 1982) and Goldberger (1983)]. This is more than a technical problem, and it is straightforward to construct realistic examples with heteros-kedasticity where the maximum likelihood estimates are worse than OLS. Particularly in survey data, where heteroskedasticity is endemic, there is no reason to suppose that Tobit will give estimates that are any better than OLS ignoring the censoring. With heteroskedasticity and censoring, neither technique is likely to give satisfactory estimates.

There are two approaches that make sense in practical applications. The first is to abandon this way of thinking about the problem. The standard approach starts from a linear model, and then complicates it to allow for censoring, treating the linearity as a maintained structural hypothesis. In the standard linear regression, this makes sense, because the structural regression coincides with the regression function, and is readily recovered from the data. In many cases, this structural assumption of linearity is merely a convenience, and there is no particular reason to believe that the underlying relationship is genuinely linear. When this is so, the standard procedure for dealing with censoring is not an attractive one, because the original linearity assumption has nothing to support it but convenience, and the convenience is lost with the censoring. The regression function (26) is not a convenient object to handle, and a more suitable alternative would be to start, not from the structure, but from some suitable direct specification for the regression function. Given the presence of the zeros, linearity might not be plausible, but some other flexible functional form might do perfectly well. I shall discuss one particular non-parametric procedure in Section 2.3 below.

The second approach is less radical, and makes sense when there is good reason to retain the linear structure. In this case, it is desirable to use an