effects of a policy or a project, we must take into account what determines it, and having done so, we will usually find that we cannot discover its effects by standard regressions. The basic issue here is the correlation of explanatory variables with the error term, and it matters less whether we think of that correlation as coming from simultaneity, heterogeneity, selection, or omitted variables.

The fully general model (36) through (38) can be estimated as it stands by using maximum likelihood once some joint distribution — typically joint normality — is specified for the three error terms u_O, u_1, and u_3. Given normality, the special case of generalized Tobit can be estimated using a short-cut technique that avoids the need for maximizing a custom built likelihood function. In a famous paper, Heckman (1976) proposed what has come to be known as the "Heckit" or Heckman's probit, by analogy with Tobit or Tobin's probit. At the first stage, the y-parameters in (37) are estimated up to scale by probit applied to a dichotomous variable that is 0 when y is censored and I otherwise. The results are then used to calculate the A-function in (40), which under normality takes the form of a Mill's ratio, which can then be used on the right hand side of (40) to estimate the /3's. This technique is very widely used in the applied development literature, although (notably) not in the study of wage equations among Panamanian males by Heckman and Hotz (1986).

The role of the distributional assumptions in these models has come under increased scrutiny in recent years. As we have already seen, maximum likelihood estimation of the Tobit model is inconsistent when homoskedasticity fails. In the general model, even identification can hinge on the distributional assumptions on the error terms, a situation that is practically little different from lack of identification altogether. The identification of the general model under minimal distributional assumptions has been addressed in papers by Manski (1988), Chamberlain (1986) and Heckman (1990). The identification of the switching equation (38) is straightforward, provided of course that we normalize the variance to unity. The identification of the structural equations in the absence of knowledge of the joint distribution of u_0, a_1, and u_2 requires that there is at least one variable in the switching equation (37) that is absent from the structural equations, although this in itself is not sufficient; for example, at least one of the variables unique to the switching equation must be continuous.

Finding variables that affect switching hut arc absent from the structure is closely akin to the general problem of finding instruments, and is frequently as difficult. In the paper that introduced selection effects into applied econometrics, Gronau (1973) found that women's wages were systematically higher when they had small children. The implausibility of children directly increasing labor market productivity led to a model in which children acted as selection