is a regression model such as (20), and that the residuals are positively correlated across observations in the same cluster. As is the case with heteroskedasticity, OLS remains unbiased and consistent, but is inefficient, and as with heteroskedasticity, our main concern is with the invalidity of standard formulas for the variance-covariance matrix of the OLS estimator. A useful result, due to Scott and Holt (1982), is that the Kish design effect is the *maximal* correction that is required, and that, in general, the estimated variances will understate the true variances by a factor that is less than the design effect. However, the maximum is attained when all the right hand side variables in the regression are constant within clusters, as would be the case when the x's are cluster prices, wages, or variables measuring access to schools, health clinics or the like [see also Kloek (1981)]. If some *x's* vary across members of the cluster, and are correlated between clusters with the other variables, the design effect will overstate the correction.

As with heteroskedasticity, there are parametric and non-parametric methods for correcting the variance-covariance matrix. Among the former would be to specify a variance components model at the cluster level, the estimation of which would allow the calculation of the intracluster correlation coefficient, which can then be used to calculate standard errors. Alternatively, an intracluster correlation coefficient can be calculated from the OLS residuals using (17) and the result used to estimate the correct variance covariance matrix for the OLS estimator. More generally, it is possible to allow for cluster fixed effects, and to work with deviations from village means. This is a useful technique in some contexts, and I shall discuss it below, but note that it does not permit us to estimate coefficients for any regressors that do not vary within the clusters.

A useful procedure is based on the fact that cluster sizes are typically small relative to the total sample size, say 10 or 16 households per cluster, so that it is possible to correct the variance covariance matrix non-parametrically by using the OLS residuals to "estimate" the variance-covariance matrix of the residuals in each cluster, just as the squared OLS residuals are used to "estimate" the variances in the heteroskedasticity-robust calculations. (I use the inverted commas around "estimate" because in neither case are we trying to obtain a consistent estimate of the individual residual variance or individual cluster residual variance covariance matrix.)

Suppose then that we have estimated the regression by OLS, and that for cluster *c* we have obtained the OLS residuals e_i. We then calculate a robust OLS variance covariance matrix by calculating [see White (1984)],

$$.1\text{-}7(i3)=(X'X)^{-1}\sum_{r} xr_c e_c e x,(X'x)^{-1} \qquad (22)$$

where X_i is the submatrix of X corresponding to cluster *c,* and C is the total