

autoregressive structures, as, for example, in Guilkey and Schmidt (1973) and Anderson and Blundell (1982). But provided autocorrelation is handled in a way that respects the singularity (as it should be), so that the omitted equation is not implicitly treated differently from the others, then it will always be correct to estimate by dropping one equation since all the relevant information is contained in the other ($n - 1$).

2.3. Estimation

For estimation purposes, rewrite (32) in the form

$$Y = \beta_0 + \sum_{i=1}^{n-1} \beta_i x_{it} + u_{it}, \quad (46)$$

with $t = 1, \dots, T$ indexing observations and $i = 1, \dots, (n - 1)$ indexing goods. I shall discuss only the case where u_{it} are independently and identically distributed as multivariate normal with zero mean and nonsingular covariance matrix Ω . [For other specifications, see, e.g. Woodland (1979)]. Since 0 is not indexed on t , homoskedasticity is being assumed; this is always more likely to hold if the β_i are the budget shares of the goods, not quantities or expenditures. Using budget shares as dependent variables also ensures that the R^2 statistics mean something. Predicting better than $w_{it} = a_i$ is an achievement (albeit a mild one), while with quantities or expenditures, R^2 tend to be extremely high no matter how poor the model.

Given the variance-covariance matrix Ω , typical element ω_{ij} , the MLE's of β_i , β_j say, satisfy the first-order conditions, for all i ,

$$E \left[\sum_{t=1}^T \frac{\partial \ln L}{\partial \beta_i} \right] = 0, \quad (47)$$

where S_{ij} is the (i, j) th element of Ω^{-1} . These equations also define the linear or non-linear GLS estimator. Since Ω is usually unknown, it can be replaced by its maximum likelihood estimator,

$$\hat{\Omega} = \left(\sum_{t=1}^T \hat{u}_{it} \hat{u}_{jt} \right) / (T - 1). \quad (48)$$

If \hat{u}_{it} replaces u_{it} in (47) and (47) and (48) are solved simultaneously, and $\hat{\beta}_i$ are the full-information maximum likelihood estimators (FIML). Alternatively, some consistent estimator of Ω can be used in place of $\hat{\Omega}$ in (48) and the resulting $\hat{\beta}_i$ used in (47); the resulting estimates of β_i will be asymptotically equivalent to FIML. Zellner's (1962) seemingly unrelated regression technique falls in this class,

