surveys unless we know that the reporting periods are the same. Exactly the same point arises if we attempt to compare two countries one of which has a perishable staple that is bought frequently, while the other uses a storable staple that is bought rarely. Problems over reporting periods and over the definition of the household are two (of the many) reasons why we know so little about international comparisons of inequality and poverty; for others [see Berry (1985) and Fields (1992)].

1.1_6. Estimation of means in stratified samples

When different households have different probabilities of being included in the survey, unweighted sample means will generally be biased for the population means. Consider the simplest example where there are two sectors, sector 1, "urban" and sector 2, "rural", and where households in each are sampled with probabilities rr_i and v·,. We are interested in the random variable *x,* which is distributed in the populations of the two sectors with means and *te,,*. There are *n* observations in total, *n,* urban households and n_2 = *n* n, rural households; these correspond to population figures of *N, N,* and N_2, so that

s =1, 2. The sample mean is *(4)*

$(ni + n2)^{-1}$ E *x,* with expectation

$$E00 \quad \overset{n_z}{+ tt2} \bullet \tag{5}$$

The population mean, *by* contrast, is given by

$$\overset{N,}{70^{21}} + \overset{N2}{711} \tag{6}$$

so that the sample mean is biased unless either $7r_1$ = ar_2, in which case the sample is a simple random sample, or $/L_i$ =- μ_2, so that the population is homogeneous, at least as far as the parameter of interest is concerned.

Neither of these requirements would usually be met in practice; for example, rural households are likely to be both poorer and costlier to sample. To get the right answer, we do the obvious thing, and compute a weighted mean. This can be done by defining "inflation factors" for each observation, equal to the reciprocals of the sampling probabilities, so that here

$$iEs,s=1,2. \tag{7}$$